

th Annual South-East European Doctoral Student Conference

Infusing Research and Knowledge in South-East Europe

Edited by: F. Gonidis, P. Gkasis, L. Lazouras, I. Stamatopoulou



Proceedings of the 8th Annual South-East European Doctoral Student Conference

Edited by

F. Gonidis, P. Gkasis, L. Lazouras, I. Stamatopoulou

ISBN 978-960-9416-06-1 ISSN 1791-3578

©SOUTH-EAST EUROPEAN RESEARCH CENTRE, 2013

24 P. Koromila str. 54622 Thessaloniki, Greece Tel: +30 2310 253 477 Fax: +30 2310 234 205 http://www.seerc.org

Preface

These proceedings represent the work of contributions to the eighth (8th) Annual SEERC Doctoral Student Conference (DSC2013) hosted by The University of Sheffield International Faculty, CITY College and organised by the South East European Research Centre - Thessaloniki, Greece. DSC 2013 has grown and continues to evolve. The Event is strongly supported by our UoS, Management School, Computer Science and Psychology departments as well as from other several departments of our university and also from the Faculty of Economics of the University of Belgrade-Serbia.

Now is the eighth year the key aim remains to provide an opportunity for PhD students to receive advice from experts in their chosen field of research. Having identified academic isolation as a problem that many doctoral students face today, SEERC aims to bring researchers together for establishing collaborative links between disciplines, for testing the ground for innovative ideas and for engaging the wider academic community.

Building on the success of the past seven conferences, this year's conference attracted a large number of submissions resulting in presentations of full papers. The audience of the conference expanded beyond the boundaries of South East Europe confirming the need for Doctoral Students to come together, discuss their experiences and gain external feedback to their work as well as listen to the progress and methodology of fellow PhD candidates. Furthermore this conference constitutes an excellent opportunity for students to present innovative ideas and practice their presentation and communication skills in front of a broad audience. Through our collaboration we can share knowledge, experiences and sometimes, even resources.

September 2013

Prof. Panayiotis H. Ketikidis Chairperson of DSC2013

Conference Organisation

DSC2013 is organized by the South-East European Research Centre (SEERC), an overseas research centre of The University of Sheffield, established as a non-profit legal entity in Thessaloniki, Greece. SEERC was founded by CITY College, the University's Intentional Faculty, in 2003.

Organising Committee

Ms. Athanasia Halatzouka, South-East European Research Centre
Dr. Lambros Lazuras, The University of Sheffield International Faculty, CITY College
Dr. Ioanna Stamatopoulou, The University of Sheffield International Faculty, CITY College
Dr. Pavlos Gkasis, The University of Sheffield International Faculty, CITY College
Mr. Fotis Gonidis, South-East European Research Centre
Ms. Danai Triantafylli, South-East European Research Centre
Mr. George Pavlidis, South-East European Research Centre
Ms. Flaka Ismaili, South-East European Research Centre

Scientific Committee Chairs

Enterprise, Innovation and Development

Dr. Pavlos Gkasis, The University of Sheffield International Faculty, CITY College Dr. Tim Vorley, The University of Sheffield

Information and Communication Technologies

Dr. Ioanna Stamatopoulou, The University of Sheffield International Faculty, CITY College

Society and Human Development: Psychology, Politics, Sociology, and Education

Dr. Lambros Lazuras, South-East European Research Centre

Scientific Committee

Enterprise, Innovation and Development

Dr. Alok Choudhary, The University of Sheffield Prof. Paul Latreille, The University of Sheffield Mr. Fotis Missopoulos, The University of Sheffield International Faculty, CITY College Dr. Nick Williams, The University of Sheffield Prof. Konstantinos Katrakylidis, Aristotle University of Thessaloniki Dr. Evangelia Desli, Aristotle University of Thessaloniki Dr. George Androulakis, University of Patras Dr. Nikos Varsakelis, Aristotle University of Thessaloniki Dr. Persefoni Tsaliki, Aristotle University of Thessaloniki Dr. Robert Wapshott, The University of Sheffield Dr. Kostas Alexandris, Aristotle University of Thessaloniki Dr. Anastasios Karamanos, The University of Sheffield International Faculty, CITY College Dr. Anna Watson, London Fashion College, University of London Dr. Howard Viney, London Business School and Open University Dr. Alexandros Kapoulas, The University of Sheffield International Faculty, CITY College Dr. Nikolaos Mylonidis, University of Ioannina Dr. Eftychia Nikolaidou, The University of Sheffield International Faculty, CITY College Dr. Leslie T. Szamosi, The University of Sheffield International Faculty, CITY College Prof. Geoff Wood, The University of Sheffield Dr. Konstantinos Priporas, Middlesex University London Dr. Tim Vorley, The University of Sheffield Prof. Christos Kollias, University of Thessaly Dr. Alexandros Psychogios, The University of Sheffield International Faculty, CITY College

Dr. Alexandros Zangelidis, University of Aberdeen

Information and Communication Technologies

Dr. Fabrizio Granelli, University of Trento,
Prof. Tudor Balanescu, University of Pitesti
Prof. Omer Faruk Bay, Gazi University
Assist Prof. Mathias Funk, Eindhoven University of Technology
Assist Prof. Alexander Chatzigeorgiou, University of Macedonia
Dr. Tony Cowling, The University of Sheffield
Dr. Konstantinos Dimopoulos, The University of Sheffield International Faculty, CITY College

- Prof. Christos Douligeris, University of Piraeus
- Dr. Dimitris Dranidis, The University of Sheffield International Faculty, CITY College
- Dr. George Eleftherakis, The University of Sheffield International Faculty, CITY College
- Prof. Neki Frasheri, Polytechnic University of Tirana
- Assoc. Prof. Akli Fundo, Polytechnic University of Tirana
- Dr. Christos Georgiadis, University of Macedonia
- Dr. Marian Gheorghe, the University of Sheffield
- Prof. Florin Gorunescu, University of Craiova
- Mr. Thanos Hatziapostolou, The University of Sheffield International Faculty, CITY College
- Prof. Florentin Ipate, University of Pitesti
- Prof. Petros Kefalas, The University of Sheffield International Faculty, CITY College
- Prof. Evangelos Kehris, Technologal Education Institute (T.E.I.) of Serres

Dr. Thomas Lagkas, The University of Sheffield International Faculty, CITY College

- Prof. Yannis Manolopoulos, Aristotle University of Thessaloniki
- Dr. Iraklis Paraskakis, South-East European Research Centre
- Dr. Ilias Sakellariou, University of Macedonia
- Prof. Demetrios Sampson, University of Piraeus
- Dr. Anthony Simons, The University of Sheffield
- Mrs. Anna Sotiriadou, The University of Sheffield International Faculty, CITY College
- Dr. Ioanna Stamatopoulou, The University of Sheffield International Faculty, CITY College
- Dr. Anca Vasilescu, University of Brasov
- Dr. Husnu Yenigun, Sabanci University
- Prof. Betim Çiço, Polytechnic University of Tirana

Society and Human Development: Psychology, Politics, Sociology, and Education

- Dr. Anastasios Stalikas, Panteion University
- Dr. Nikolaos Tsigilis, Aristotle University of Thessaloniki
- Dr. Tom Stafford, The University of Sheffield
- Dr. Chris Jones, The University of Sheffield
- Dr. Elvira Masoura, Aristotle University of Thessaloniki
- Dr. Mary Kosmidis, Aristotle University of Thessaloniki
- Dr. Antonia Ypsilanti, Aristotle University of Thessaloniki
- Dr. Angeles F Estevez, Universidad de Almeria
- Dr. Kwang-Hyuk Lee, The University of Sheffield
- Dr. Panagiotis Bamidis, Aristotle University of Thessaloniki
- Dr. Suzie Savvidou, The University of Sheffield International Faculty, CITY

College

- Dr. Elissavet Chrysochoou, The University of Sheffield International Faculty, CITY College
- Dr. Martha Kelpi, The University of Sheffield International Faculty, CITY College

Dr. Lambros Lazuras, The University of Sheffield International Faculty, CITY College

Dr. Tatiana Tairi, Massey University

Dr. Efrosini Kalyva, The University of Sheffield International Faculty, CITY College

Ms. Agnes Brunet, The University of Sheffield International Faculty, CITY College

Dr. Sara Hannam, Oxford Brookes University

Dr. Nicola Rehling, Aristotle University of Thessaloniki

Dr. Paul Stubbs, Institute of Economics Zagreb

- Dr. Natalia Vlas, Babes-Bolyai University
- Prof. Vasile Boari, Babes-Bolyai University

Prof. Theodore Chadjipandelis, Aristotle University of Thessaloniki

- Dr. Christos Frangonikolopoulos, Aristotle University of Thessaloniki
- Prof. Ralph Negrine, The University of Sheffield

Dr. Pavlos Kotsidis, The University of Sheffield International Faculty, CITY College

Dr. Filipos Proedrou, The University of Sheffield International Faculty, CITY College

Dr. Zoi Tatsioka, The University of Sheffield International Faculty, CITY College

Dr. Cormac Behan, The University of Sheffield

Dr. Alexandra Kaasch, The University of Sheffield

Sponsoring Institutions

South-East European Research Centre (SEERC), Greece

The University of Sheffield International Faculty, CITY College, Greece

The University of Sheffield, United Kingdom

University of Belgrade, Serbia

Table of Contents

Enterprise, Innovation and Development

Opportunities of increasing livestock production efficiency by developing agro-services in Uzbekistan	19
CRM performance optimization and customer base augmentation through innovation and building strong affective commitment with customers vis-a-vis behavioral commitment	32
Dynamic of Social-Economic disparities and regional development in Albania	50
What Factors do Investments Depend on —The Case of Serbia Dragana Draganac	64
Proactive Feedback-Seeking Behaviour as a Valuable Resource for Nascent Social Entrepreneurs: Entrepreneurial Opportunity Evaluation and Refinement	78
Investigating the causality between Financial Development and Economic Growth in the developing countries of Europe: Evidence from Albania and Turkey <i>Esida Gila Gourgoura and Eftychia Nikolaidou</i>	90
Key properties of the enterprise-universities interorganizational networks and the capacity for introducing innovations	107
The Role of Networking on SMEs Innovation: the case of Kosovo Flaka Ismaili	113
Examination of Competition Policy Effects on the Retail Market Development in Transition Countries by Applying the Panel Model Sinisa Milosevic	123
Building economic institutions and reforming the economy: the case of Kosovo Linda Tahiri Rukiqi	139

Comprehending HRM policies and practices in multinationals within the hospitality sector: country of origin and country of domicile effects <i>Giovanni Oscar Serafini</i>	145
Family Business in Greece. Investigation of the basic problems Sotirios Vlachakis, Marja Naaranoja, and Kerstin Siakas.	172
The journey of business location factors through years: a literature review Harissia Vlachou, and Olga Iakovidou	188
Information and Communication Technologies	
Using PKI to increase the security of the electronic transactions Marin Aranitasi, Marsida Ibro, and Bojken Shehu	224
Simulation of Network Request Mapping Techniques Adelina Basholli, and Thomas Lagkas	233
On the Role of Stream Reasoning in Run-time Monitoring and Analysis in Autonomic Systems	247
Research Method For Studying Complex Phenomenon In Information Systems Ehimen Ejodame, and Stuart Maguire	259
Existing approaches for cross-platform development and deployment of cloud applications Fotis Gonidis, Iraklis Paraskakis, and Anthony J. H. Simons	270
High-Throughput Asymmetric Encryption cryptanalysis in Embedded Systems	275
Semantic Web Trends on Reasoning Over Sensor Data Edmond Jajaga, Lule Ahmedi, and Lejla Abazi Bexheti	284
Enrichment of Digital Libraries with Web 2.0 Resources for Enhanced User Search Experience	294
Towards Continuous Quality Assurance in Future Enterprise Cloud Service Brokers Dimitrios Kourtesis, and Konstantinos Bratanis	301
Cloud Service Portability Opportunities Magdalena Kostoska, Marjan Gusev, Sasko Ristov, and Kiril Kiroski	313

Performing a Penetration Test in a Data Network Renalda Kushe, and Adrian Shehu	321
Green Cloud Computing Trends: A Review Enida Sheme, and Neki Frasheri	331
Improved MDS-MAP Algorithm for Nodes Localization in 3D Wireless Sensor Networks	340
Wireless Sensor Networks Framework for Indoor Temperature Regulation Biljana Stojkoska and, Andrijana Popovska Avramova	354
Validation Study for Two-Stage Vessel Segmentation Algorithm Radu Teodoru Popa, and Mihai Mocanu	365
Society and Human Development Psychology, Politcs, Sociology and Education	
The relationship between dysfunctional patterns of families response to the illness and symptom severity in adolescent patients with anorexia nervosa at illness onset: A gender-specific approach Dimitra Anastasiadou, Ana Rosa Sepulveda, and Montserrat Graell	377
Reading Literacy in the Educational Process from Primary School Teachers Viewpoint Jana Drabova, Tereza Gajdokova, and Vlasta Rerichova	386
Developing a Workplace Cyberbullying Measure Sam Farley, Carolyn Axtell, Christine Sprigg and Iain Coyne.	399
Exploring the Gesamtkunstwerk concept in the design of the theatrical spatial experience. A constructivist approach	404
Mapping the obstacles to effective governmental and non-governmental action against drug trafficking: the sociopolitical context in Serbia Darja Koturovic	413
Including female non mainstream representations in NGO mainstream discourses to enhance peace building	421
Greek memorandums of understanding and democracy Alexandros Kyriakidis	433
Children's Healthcare Facilities: Architectural aspects for the improvement of the environments quality Artemis Kyrkou, and Fani Vavili	467

A systematic literature review on the correlations of cognition and everyday functionality assessment tools for healthy older adults <i>George Pavlidis, and Ana Vivas</i>	476
The development of executive functions in children Fatbardha Qehaja-Osmani	489
Patterns of Stressors Among Albanian University Students: The Case of the European University of Tirana Rudina Shkullaku and Erika Melonashi	495
The importance of children's participation in decision-making about themselves and the barriers, which might encounter	506
The Role of Optimism Bias in Greek Young Drivers Propensity to Commit Road Violations Danai Triantafylli	516
Author Index	527

Enterprise, Innovation and Development

Opportunities of increasing livestock production efficiency by developing agro-services in Uzbekistan

Akmal Abruev¹, Olim Murtazaev¹ and Stamatis Aggelopoulos²

¹(Agricultural Economics and Management Department) Samarkand Agricultural Institute, Samarkand, 140103, Uzbekistan ¹samai_akmal@yahoo.com ²(Agricultural Economics and Management Department) Samarkand Agricultural Institute, Samarkand, 140103, Uzbekistan ²olimmurtazaev@gmail.com ³(Agricultural Development & Agribusiness Management Department) Alexander Technological Educational Institute (ATEI) of Thessaloniki, Greece ³stamagg@farm.teithe.gr

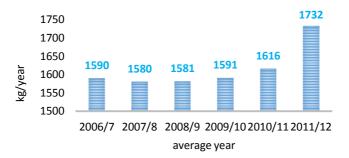
Abstract. The object of the work is to study the level of inter-relatedness of efficiency of livestock products and service branches. Analyzed object is studied in the case of private farms, which engaged in dairy production, and agricultural service branches. 114 private farms surveyed as an optional way. Likewise, we used secondary data concerning to livestock breeding and raising of Uzbekistan, state committee of statistics, scientific articles, reports and internet sites. We analyzed influence of multiple correlation and regression with SPSS software. According to obtained results, there is strong correlated relations between milk yield and fodder service, AI service and veterinary service. Fodder service is marked as statistically significant among service branches. In the future, for increasing milk yield in Uzbekistan we should develop fodder service, AI service and veterinary service extensively, that is via increasing quality of services and introducing innovation technology news.

Keywords: Agro-service, Development, Efficiency, Uzbekistan

1. Introduction

Result of production in agro industry is up to agriculture but also the developed level of service branches (Ismoilov & Murtzaev, 2005). Quality of service is marked with production efficiency. Service branch lines is increased drastically in the last years in Uzbekistan (Appendix 1.). The number of livestock heads for last decade increased too (Appendix 2.). The policy act carried out in agricultural sector of national economy have effected a gradual increase in production - both crop and livestock (Mukhitdinova, 2010). Nowadays, there are households (dehkan), private farms and agricultural cooperatives (shirkat) engaged in livestock production (Appendix 3.). In all categories of economy the number of cattle has constituted 9642,5 thousand head and has increased on 548,8 thousand heads or by 6 percent in 2012 (SCSUz, 2013).

The number of cows head has increased on 84,6 thousand heads (by 2,23 percent). The number of cattle including cows, sheep and goats, poultry has increased in all Provinces of the Republic. However, the increase of livestock head doesn't mean the increase of efficiency in the sector. Average milk yield has constituted 1732 kg/year during 2011/12 (Fig.1). Lerman's researches indicated that milk yield was 1500 kg/year in 1980, 1600 kg/year in 2000/2008 in Uzbekistan. This indicator is low even according to indices of the European countries, USA, and numerous of CIS countries. In sum, Bulgaria 3738 kg/year, Romania 3822 kg/year, Slovenia 5451 kg/year, Slovakia 5642 kg/year, Italia 6739 kg/year, Austria 6717 kg/year, Hungary 6846 kg/year (EC, 2011), US 9681,9 kg/year (USDA, 2012), Kazakhstan 2253 kg/year (Haas, 2011), Ukraine 4121 kg/year (Engelen, 2011), Russia 3776 kg/year (Tarassevych, 2011). In these countries livestock productivity is achieved by increasing pedigree animals, wisely usage of agro services and new technologies. Increasing of efficiency indices (milk yield, meat per head, calves per cow) in livestock breeding and raising has an impact on efficiency of farms and exceeding of competitiveness in market economy.





Last years are marked with positive changes in milk yield in the Republic. This change is due to the state attention of strengthening of breeding potential, developing livestock breeding farming societies, increasing of pedigree cattle in them, developing and widening of breeding nucleus, artificial insemination of cows and importing pedigree cattle from foreign countries. (DPRU_308, 2006; DPRU_842, 2008). It is important to take care of pedigree cattle and delivering necessary top quality agro services on time. One of the main objectives of the work is to study present state of these services and their usage of development in near future.

The purpose of this research is to study the level of correlation of efficiency of livestock breeding and raising products with agricultural service branch lines. Research questions: Do the service branches provide efficient services? Which services should be developed for increasing of livestock production efficiency? Based on a sample 114 livestock farms, Multiple Correlation and Multiple Linear Regression are estimated, which means that holding inputs constant and suggesting farmers how they can get more output by developing and enhancing agro service branches.

2. Literature Review

Cattle beef and dairy milk are considered as rich for carbohydrate agricultural products. Their production and delivering to consumer is considered as an important. Nevertheless, amount lack of production and lack of necessary processing technologies are restricting opportunity of full coverage of markets. World Bank (2010) reported that, one of the more vulnerable categories than other European and Central Asian countries, in Uzbekistan the share of food as a percentage of total household expenditure. As well, numerous livestock products of dehkan farms are sold to neighboring houses and markets near these areas.

There are several researches have been done about agro service development and efficiency measurement of livestock sector in Uzbekistan. The analysis (Mirzaev, 2011) indicates the issues related both with traditional and institutional approach barriers hindering on the development of livestock services. Such problems as, financial resources; organization and economical character in the activity of service providers; lack of proper management and skills in dealing with marketing and prognosis of livestock market production and trade; accessibility of service providers to bank loans; and insurance system. In suggestions, offer complex of institutional and economic policy factors should be involved for developing livestock services. Another scientists (Yusupov, Lerman, Chertovitskiy, & Akbarov, 2010) views of low cattle sub-sector efficiency of Uzbekistan aggravated by lack of competitive input and service supply markets. Previous study analysis of Abruev et al., (2012) have been applied nonparametric approaches to estimate relative efficiency of livestock farms per whole rayon and concludes to greater extent welfare of farms efficiency depend on stock quality breeds and cattle management. Energy use is one of the main cost expenditure in production. Wind, hydropower and solar are capture energy flows available from the environment (Kapiki, 2010). Biogas energy are one of the energy efficiency source and large livestock farms managers should to implement to their scale operation as second source of income.

Yuldashbaev (2011) argued four the main factors as: insufficient feed resources; lack of land areas and turnover; lack of credit resources; difficult access to inputs and **services**, which negatively effecting Uzbekistan's livestock sector. Likewise, the lack of organizations of agro services in rural areas constraining the economic efficiency of manifold of agricultural producers (UNDP, 2008; Djanibekov, Lamers, & Bobojonov, 2010). Abruev and Rahmonov (2011) stated that zoo veterinary branches play a key role in increasing of livestock breeding and raising and they should be well-equipped and for this reason, to increase assigned funds to equip with necessary equipments, to develop works in artificial insemination (AI) stations and their stimulation by the state. Thus, the importance of achieving efficiency in looking after services is realise to be important.

Proper organization and implementation of agro service infrastructure that meets the world standards is becoming actual. Enhancing and development of sub- service sectors (veterinary, artificial insemination, forage, processing, banking services) remaining the main problem not only for private farms, but also in households (*dehkan*) in Uzbekistan agriculture.

3. Data and Methodology

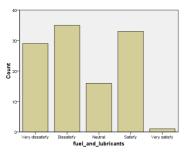
Amount of cattle in Samarkand region has constituted 1.284.951.0 in 2012 and it has a strategic importance in the development of cattle breeding of Uzbekistan. According to development tempo of the left two agricultural enterprises (shirkat) and dehkan societies are not better than private farms in the direction of livestock breeding. In addition, there are good mutual contracts between private farms and agro service branches. For that reason, services of agro service branches are closely connected with customer satisfaction survey and society's economic state, in our example this is done due to ascertaining of correlation between cows' milk yields. 114 private farms in the direction of cattle breeding or total 19,4 % of farmers surveyed in 587 cattle breeding private farmers of all the districts of Samarkand Province. Average cattle heard was 55,37 head in private farms.

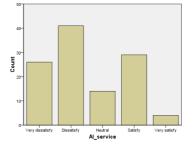
Data were collected on February-April of 2013 and previous year's (2012) condition was thoroughly studied. Managers of milk producing private farms (or vice manager as in the absence of manager) participated in the questionnaire. In separate sheet managers had opportunity lived their briefly comments as well. Questionnaires were collected separately in work places, offices of farms and fields without interrupting with producing process. It should be noted that, as it were ascertained managers of private farms participated first time in the questionnaire. It drawn that, agro service branches have not organized statistical feedback from producers until this time. This gives rise to the customer gap. The study used multiple correlation and multiple regression analysis to examine models by IBM SPSS 20.0 software.

4. Empirical Analysis

4.1. Graphs

Finding graphs:





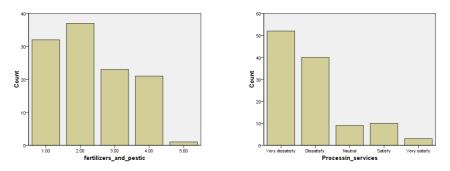


Fig. 2. Customer satisfaction

Fuel & lubricants; Fertilizers & pesticide service; and Processing services are mostly dissatisfied customer demand. Insufficient supply Fuel & lubricants; Fertilizers & pesticide service leads to low productivity and artificially high prices increase the cost of production of milk and meat in general.

Knowledge and experience of veterinarians has an important role for servicing veterinary and AI services to consumers. Treating of animals by veterinarians on time, right diagnosis for illnesses, skilled performing of treating and AI has a direct influence on livestock productivity and efficiency of farms.

It should be noted that the majority of farmers do not caring for the environment and sanitary standards, and animal house constructions are not suitable for set automatization, that bring to the difficulties in providing agro services.

4.2. Correlation Analysis

Bivariate Correlation applied to identify the correlation between variables. There are dependent variables is milk yield per cow and independent variables are agro service braches.

	Table 1/a. Test of Correlation							
	Correlations							
				INDEPENDE	ED VARL	ABLES		
		1	2	3	4	5	6	7
	ENDENT RIABLE	Machine & tractor service	Water users service	Veterinary service	AI service	Vet drugs service	Fodder service	Fuel & lubricants service
Milk yield	Pearson Correlation	.281**	.335**	.718**	.686**	.521**	.678**	.304**
per cow	Sig. (2- tailed)	.002	.000	.000	.000	.000	.000	.001
	Ν	114	114	114	114	114	114	114

Table 1/a Test of Correlation

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

			Table 1/b. Test of Correlation					
Correlations								
		Π	NDEPENDEI) VARIA	ABLES			
	8	9	10	11	12	13	14	
ENDENT IABLE	Fertilizers & pesticide service	Bank service	Consulting service	Prod. sale serv.	Transport service	packing	sing	
Pearson Correlation	.466**	.508**	.616**	.597**	.335**	.432**	.352**	
Sig. (2- tailed)	.000	.000	.000	.000	.000	.000	.000	
N	114	114	114	114	114	114	114	
	IABLE Pearson Correlation Sig. (2- tailed)	NDENT IABLE& pesticide servicePearson Correlation.466**Sig. (2- tailed).000	89NDENT IABLEFertilizers & pesticide serviceBank & servicePearson Correlation.466**.508**Sig. (2- tailed).000.000	INDEPENDEL8910FertilizersBankConsultingNDENT IABLE& pesticide serviceserviceservicePearson Correlation.466**.508**.616**Sig. (2- tailed).000.000.000	INDEPENDED VARIA891011FertilizersBankConsultingProd.NDENT IABLEServiceserviceservicesale serv.Pearson Correlation.466**.508**.616**.597**Sig. (2- tailed).000.000.000.000	INDEPENDED VARIABLES89101112FertilizersBankConsultingProd.TransportNDENT IABLE\$esticide serviceservice servicesale serviceservice servicePearson Correlation.466**.508**.616**.597**.335**Sig. (2- tailed).000.000.000.000.000	INDEPENDED VARIABLES8910111213NDENT IABLEFertilizers erviceBank serviceConsulting serviceProd. sale serviceTransport packing serviceTare & packing servicePearson Correlation.466**.508**.616**.597**.335**.432**Sig. (2- tailed).000.000.000.000.000.000	

Table 1/b Test of Correlation

The correlation coefficient of milk yield strong relationship with fodder service, AI service, and veterinary service, has constituted 678^{**}; .686^{**}; and .718^{**} respectively. Correlation of factors influencing indirectly to changing of milk efficiency was 2tailed**. Machine & tractor service; Water users' service; Fuel & lubricants; Transport services; processing services were between 0.2≤4 and weak relationship was observed. Many farmers have their own tractor fleet, and prefer to use them, despite the high cost of maintaining machine & tractor parking service. Because, they are independent and have an opportunity to use them at any time of the season. In addition, a large number of farmers have their own wells for irrigation as delivered by the pump or irrigation canal. According to this, the correlation between farms and these two independent variables has very weak relationship.

4.3. Output of Multiple Regression Analysis

According to results of multiple correlation analysis, we analyzed directly influencing three variables to milk yield in multiple regression analysis that is fodder service, AI service and veterinary service.

Model Summary									
Model	R	R Square	Adjusted R	Std. Error of the Estimate					
Square									
1	.826 ^a	.682	.638	339.32593					
a. Predictors: (Constant), Processing services, Machinery & tractor parking service, Fuel &									
lubricants service, Production sale service, Water users service, Fertilizers & pesticide									
service, Tra	ansport services	, Fodder servi	ces, Vet drugs serv	vice, Tare & packing service, Bank					

Table 2. Summary statistics for survey variables	Table 2.	Summary	statistics	for	survey	variables
--	----------	---------	------------	-----	--------	-----------

service, AI service, Consulting service, Veterinary service R is equal .826 value, indicates a good level of prediction. R^2 independent variables are 68.2% of the variability of milk yield.

ANOVA ^a								
Model		Sum of Squares	df	Mean Square	F	Sig. .000 ^b		
1	Regression	24502732.959	14	1750195.211	15.200	.000 ^b		
1	Residual	11399066.795	99	115142.089				

Table 3. ANOVA

Total	35901799.754	113
-------	--------------	-----

a. Dependent Variable: Milk yield per cow

b. Predictors: (Constant), Processing services, Machinery & tractor parking service, Fuel & lubricants service, Production sale service, Water users service, Fertilizers & pesticide service, Transport services, Fodder services, Vet drugs service, Tare & packing service, Bank service, AI service, Consulting service, Veterinary service

The general form of the equation to predict milk yield from directly effecting variables such as fodder service, AI service, and veterinary services. Findings of (Lerman, 2008) was explored more rigorously by regression analysis has a statistically significant positive effect on milk yields: of feed sufficiency, human capital and the use of artificial insemination by livestock farms.

	(Coefficients ^a			
Model	Unstandard	ized Coefficients	Standardized Coefficients	t	Sig.
	В	Std. Error	Beta		
(Constant)	55.736	164.997		.338	.736
Machinery& tractor service	35.999	47.366	.050	.760	.449
Water users	19.705	39.600	.035	.498	.620
Veterinary service	100.723	55.117	.205	1.827	.071
AI service	71.131	48.289	.151	1.473	.144
Vet drugs service	8.442	38.172	.018	.221	.825
Fodder service	127.768	41.305	.251	3.093	.003
Fuel and lubricants	-10.755	33.814	023	318	.751
Fertilizers and pesticide	19.178	40.285	.037	.476	.635
Bank service	27.226	43.623	.049	.624	.534
Consulting service	37.187	42.681	.078	.871	.386
Production sale service	94.071	40.528	.189	2.321	.022
Transport services	7.359	36.225	.015	.203	.839
Tare & packing service	2.741	46.782	.005	.059	.953
Processing service	35.448	39.269	.066	.903	.369

a. Dependent Variable: Milk yield per cow

A multiple regression was run to predict milk yield from directly and indirectly effecting services. These variables statistically significantly predicted milk yield, F(14, 99) = 15.200, p<.005, R²=.682. We can see from the Sig. column that fodder service variable added statistically significantly to the prediction, p<.005. Production sale service, veterinary service, AI service independent variables coefficients are less statistically significantly.

5. Conclusions

5.1. Conclusions and Suggestions

Mechanisms of conformation and management of service branches' activity have been developed last years. Now we should develop them according to modern international standards. By turn, this influences on increasing efficiency in livestock breeding then afterwards quality of agricultural products and export.

It was ascertained strong correlation between milk yield and agro service branches such as fodder service, AI service, and veterinary service has constituted 678**; .686**; and .718** respectively. Fodder service is marked as statistically significant 0.03 (p<0.05) among service branches. According to the results of analysis, intensive and extensive developing of fodder, AI and veterinary services in near future has an important role in increasing of milk yield. For this reason, need Government support and agricultural policy to implement innovation, development and quality management in dairy farms. For this occasion:

• To apply scientific innovation and technical achievements for instance, via Growing for sprouted fodder for livestock (Wiskerchen, 2012) we can fully supply demand for producing fodder in small farms and households; Enlarging existing plots of land for increasing cowherd's farms. Government support to dairy farms, with free access to purchase fertilizers, other chemicals, fuel-oiling materials.

• To organize fodder stores with mixed fodder, cottonseed meal, husk selling unions in the centers of districts, and conclude contracts with farmers for small-scale delivery. According to demand of most consumers, their bought services should be delivered on time and top quality conditions.

• Increase amount of preferential long-term credits (loans) to innovate animal house and implement AMS (automatic milking system) in dairy farms. Transform to collect the milk in milk -tanks rather than in smaller (unfreeze) containers.

• Veterinary association and agricultural institutes should organize seminarcourses of improvement of professional skills and exchange of experience. Also, consulting services that influencing indirectly on producing efficiency is important in strengthening experiences in newly established private farms and delivering of innovations on time.

Wisely, usage of all kinds of opportunities of increasing efficiency livestock first directly connected with amount of fodder resources, quality and composition. (Ismoilov & Murtzaev, 2005). Important source of increasing country's product fund is via servicing manufacturer on time and quality agro services, ending of agricultural product loss and improvement of usage. Because in return for not obeying rules of producing, keeping, loading unloading, also owing to spoiling of product quality every year damaged as million sum¹ s. By developing of agro service and strengthening, we can achieve positive results in livestock sector Yuldashbaev (2011). Extension service, veterinary service, artificial insemination service, forage making service, selection and seeding services are important for this reason. Mirzaev (2012)

¹ Uzbek national currency 1.000 sum = $0.4 \in$.

stated differentiation method of service prices in developing economic relations of agro services market has a particular importance. To apply such services period terms and execution of work and services development.

Further development of agricultural production and the service is closely related to the implementation of innovation to services. Necessary to attract private and public investment to improve the quality of agricultural services. Based on the small size and operation length of farms, households - innovation can be implemented through services branches.

5.2. Limitation and Recommendations for future research

Ties between service branches of agriculture and production efficiency indices in livestock breeding and raising were studied. It desired to compare financial results of private farms. However, selected theme is too important and it should be well discussed, this presented article may be a good basis for the future researches.

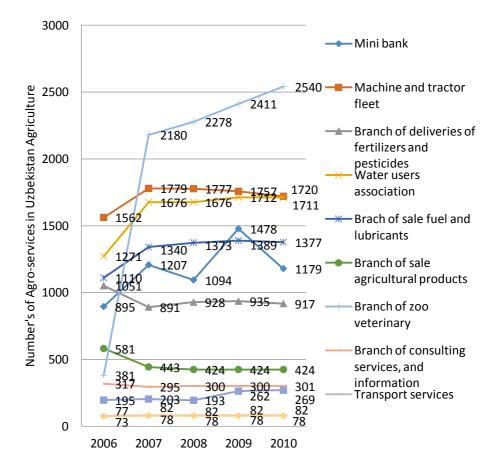
References

- Abruev, A., & Rahmonov, D. (2011) Some issues on improving zoological veterinary services. *University of Oradea CNCSIS B+journal*, X/B, p.263-268.
- Abruev, A., Theocharopoulos, A., & Aggelopoulos, S. (2012) Technical efficiency and management of livestock production: the case of republic of Uzbekistan. 2nd Advances in Hospitality & Tourism Marketing and Management Conference (AHTMMC). Corfu Island, p.1-7.
- Djanibekov, N., Lamers, J., & Bobojonov, I. (2010) Land Consolidation for Increasing Cotton Production in Uzbekistan: Also Adequate for Triggering Rural Development, in: Labar, K., Petrick, M. and G. Buchenrieder (eds.). Challenges of education and innovation - Proceedings of the Fourth Green Week Scientific Conference. IAMO: Studies on the Agricultural and Food Sector in Central and Eastern Europe. Halle (Saale), p.140-149.
- DPRU_308, (2006) Degree of the President of the Republic of Uzbekistan "On measures of breeding livestock and widening of producing livestock products in private helpers, dehkan and farming societies". Collection of legislation of the Republic of Uzbekistan, Tashkent
- DPRU_842, (2008) Degree of the President of the Republic of Uzbekistan "On additional measures of breeding livestock and widening of producing livestock products in private helpers, dehkan and farming societies". Collection of legislation of the Republic of Uzbekistan, Tashkent
- EC, (2011) EU Dairy farms. European Communities, Brussels
- Engelen, A. (2011) Dairy Development In Kazakhstan. Food And Agriculture Organization Of The United Nations, Rome

- Francksen, T., Hagemann, M., & Latacz-Lohmann, U. (2012) Growth of milk production in German dairy farms: an empirical study based on event history analysis. *Agricultural Economics*, 43(6), p.671-685.
- Haas, M., & Maksimenko, M. (2011) Russian Federation Dairy and Products Annual. USDA Foreign Agricultural Service - Global Agricultural Information Network, Washington D.C.
- Ismoilov, A., & Murtazaev, O. (2005) Agricultural Economics. Moliya, Tashkent
- Kapiki, S. (2010) Energy Management in Hospitality: a Study of the Thessaloniki Hotels. *Economics and organizatoin of future enterprise*. p.78-97.
- Leeper, J. (2012) *What statistical analysis should I use?* [Online]. University of California, Academic Technology Services: Available at: <u>http://www.ats.ucla.edu/stat/stata/ado/analysis/</u>. [Accessed 1 May 2013].
- Lerman, Z. (2008) Sustainable Development of the Livestock Sector in Uzbekistan: Status and Policy Recommendations. POLICY BRIEF, UNDP. 1(11), p.1-12.
- Mirzaev, K. (2011) Approaches and issues for developing livestock services in Uzbekistan Perspectives of Innovations, Economics & Business. *International Cross-Industry Journal*. 8(2), p.23-25.
- Mirzaev, K. (2012) *Methodical bases of developing efficiency agro services*. Samarkand Institute of Economics and Services, Samarkand
- Mukhitdinova, U. (2010) Potential of production and trade of agricultural products in Uzbekistan. Perspectives of Innovations, Economics & Business. *International Cross-Industry Journal*. 5(2), p.45-46.
- SCSUz. (2013). *The State Committee of the Republic of Uzbekistan on statistics*. Available at: <u>http://www.stat.uz</u> [Accessed 1 April 2013].
- Tarassevych, O. (2011) Ukraine Dairy and Products Annual. USDA Foreign Agricultural Service, Global Agricultural Information Network, Washington D.C.
- USDA (2012) *Milk Production*. National Agricultural Statistics Service (NASS), Agricultural Statistics Board, United States Department of Agriculture (USDA), Washington D.C. p.21.
- Wielkiewicz, R. (2012) SPSS online guide, [Online]. Available at: <u>http://www.mhhe.com/socscience/psychology/runyon/spss/conclusion.html</u>. [Accessed 1 May 2013].
- Wiskerchen, J. (2012) *Growing Sprouted Fodder For Livestock*, [Online]. Peakprosperity.com's ResilientLife how to prepare for a better future. Available at: <u>http://www.peakprosperity.com/</u>. [Accessed 5 May 2013].
- Yusupov, Y., Lerman, S., Chertovitskiy, A., & Akbarov, O. (2010) Animal industries in Uzbekistan: the current condition, problem and prospect of development. The analysis in a context of the tendencies of development of agrarian sector. Nasaf, Tashkent

Appendix

Appendix 1.



Graph. 1. Changing number of infrastructure objects in Republican agriculture, (SCSUz, 2006-2010)

Appendix 2.

	В	y Republic		By Samarkand province				
Year	Number	Including:	By	Number of	Including:			
S	of cattle,	cows,	total	cattle,	cows,	By total		
	thousand	thousand	cattle	thousand	thousand	cattle		
	head	head	head,	head	head	head,%		
			%					
1996	5102,5	2234,2	43,8	733,1	323,1	44,1		
1997	5196,4	2281,3	43,9	771,2	335,9	43,6		
1998	5225,2	2290,2	43,8	781,7	343,6	44,0		
1999	5281,8	2310,0	43,7	728,1	346,0	47,5		
2000	5353,4	2343,4	43,8	794,0	383,5	48,3		
2001	5416,1	2361,8	43,6	790,2	354,3	44,8		
2002	5477,6	2393,2	43,7	807,4	362,3	44,9		
2003	5878,8	2556,7	43,5	876,6	403,6	46,0		
2004	6242,7	2704,0	43,3	965,5	451,1	46,7		
2005	6571,4	2821,3	42,9	996,8	483,5	48,5		
2006	7044,6	2982,5	42,3	1006,2	498,3	49,5		
2007	7458,1	3124,9	41,9	1045,8	510,7	48,8		
2008	8024,8	3327,0	41,5	1104,4	521,4	47,2		
2009	8510,8	3535,6	41,5	1118,2	538,3	48,1		
2010	9093,7	3793,6	41,7	1191,3	570,8	47,9		
2011	9642,5	3878,2	40,2	1216,8	578,0	47,5		
2012	10141,3	3935	38.8	1284,9	588,1	45.7		

 Table 1. Changing of cattle by years (according to all kinds of societies), (SCSUz, 2011)

Appendix 3.

 Table 2. Production of main livestock breeding by categories of societies, (SCSUz, 2011)

2011)							
	2011	2010	Growth rate				
			for 2009, in				
			%				
	1564,2	1461,4	106,8				
Meat in live weight, thousand ton							
including:							
agricultural cooperative society (<i>shirkat</i>)	41,6	34,4	106,6				
private farms	41,0	37,8	106,8				
household farms	1481,6	1389,2	106,9				
Milk, thousand.ton	6766,2	6169,0	106,7				
including:							

41,5 230,1	36,2 205.0	124,8 116,1
6494,6	5927,8	106,4
3441,7	3058,8	112,6
1191,7	995,2	113,6
374,5	288,1	137,2
1875,5	1775,5	108,9
28687,0	26510	106,1
2551,0	2844	98,2
1938,0	1715	119,2
24198,0	21951	106,3
1022,3	934,9	104,1
335,3	258,7	95,6
43,4	44,1	104,2
643,6	632,1	108,1
	230,1 6494,6 3441,7 1191,7 374,5 1875,5 28687,0 2551,0 1938,0 24198,0 1022,3 335,3 43,4	$\begin{array}{ccccccc} 230,1 & 205,0 \\ 6494,6 & 5927,8 \\ 3441,7 & 3058,8 \\ \hline 1191,7 & 995,2 \\ 374,5 & 288,1 \\ 1875,5 & 1775,5 \\ 28687,0 & 26510 \\ \hline 2551,0 & 2844 \\ 1938,0 & 1715 \\ 24198,0 & 21951 \\ 1022,3 & 934,9 \\ \hline 335,3 & 258,7 \\ 43,4 & 44,1 \\ \hline \end{array}$

Appendix 4.

 Table 3. Questioner- Agro services indexing (customer satisfaction)

	Very dissatisfy	Dissatisfy	Neutral	Satisfy	Very satisfy	Briefly comments
"Machinery & tractor stations services	1	2	3	4	5	*
Water usage services	1	2	3	4	5	*
Veterinary services	1	2	3	4	5	*
AI service	1	2	3	4	5	*
Zoo vet drugs	1	2	3	4	5	*
Fodder service	1	2	3	4	5	*
Fuel and lubricants services	1	2	3	4	5	*
Fertilizer & lubricants services	1	2	3	4	5	* *
Mini bank services	1	2	3	4	5	*
Consulting services	1	2	3	4	5	*
Production services	1	2	3	4	5	*
Transport services	1	2	3	4	5	*
Tare and packing	1	2	3	4	5	*
Processing services	1	2	3	4	5	*
Average milk yield per cow (kg/year), in 2012						

CRM performance optimization and customer base augmentation through innovation and building strong affective commitment with customers vis-a-vis behavioral commitment: Comparative Overhaul to delineate the importance of innovation and established relationships with customers in order to uncover their galvanizing effects on the resilient long term customer loyalty representing a successful strategy for building sustainable competitive advantage for Multinational enterprises owners of Airports in Southeastern Europe

Ilija Cvetanovski¹ and Dr. Alexandros Kapoulas²

BAED, City College, International Faculty of the University of Sheffield ¹ilijac.email@gmail.com ²akapoulas@city.academic.gr

Abstract: The following paper will examine the research query that aims to provide understanding about an emerging phenomenon that will investigate the importance of building affective commitment with customers on the expense of behavioral commitment followed by permanent innovation, representing critical factors of a successful business strategy that lead to creation of sustainable competitive advantage for Airports enterprises owners in Southeastern Europe. Why is the topic so pertinent and considered to be global emerging phenomenon? For the reason that in the present highly competitive environment and turbulent market, the customer is more important than ever.

Considering the fact that this emerging phenomenon is unexplored, Wachnik, B. (2012), Valvi and Fragkos (2012), Hong and Wang (2012), Bulearca and Tamarjan (2012), Bhradwaj, et al (2012), Livne, *et al* (2011), Slieme, F. (2010) and Gwang-Suk, et al (2012) present the need for further research in various geographic and industry settings.

For the particular research, literature by Bulearca and Tamarjan (2010), Fitzgibbon and White (2005), Menon and O'Connor (2007) and Kapoulas, *et al* (2002) suggest that most suitable research approach is qualitative. Creswell (1994) and Kapoulas and Mitic (2012) agree that by following an interpretative research through inductive theory that generates "mode of inquiry" researchers could discover a production of substantial and valuable knowledge for the academic and business interest groups. Moreover suggestions from earlier research, semi-structured in-depth interviews, comparative case studies and focus groups will be conducted.

Keywords: *affective commitment, competitive advantage, innovation*

1. Introduction

Global economic turbulence, rapid market condition variance and rising competitive power have brought enterprises to unstable grounds with rising unpredictable variables regarding future market directions and customer preferences. The following paper will examine the research query that aims to provide understanding about an emerging phenomenon that will investigate the importance of building affective commitment with customers on the expense of behavioral commitment followed by permanent innovation, representing critical factors of a successful business strategy that lead to creation of sustainable competitive advantage for Airports enterprises owners in Southeastern Europe. Why is the topic so pertinent and considered to be global emerging phenomenon? For the reason that in the present highly competitive environment and turbulent market, the customer is more important than ever. This research will substantially contribute to the enrichment of academic research considering the fact that it investigates a global emerging phenomenon. (Kapoulas, et al, 2002)

The necessity for investigation of this emerging phenomenon derives from recommendations of earlier academic research by Wachnik, B. (2012), Valvi and Fragkos (2012), Hong and Wang (2012), Bulearca and Tamarjan (2012), Bhradwaj, et al (2012), Slieme, F. (2010), Brown, H. (2012), Livne, *et al* (2011), Menon and O'Connor (2007), Fitzgibbon and White (2005), Battor and Battor (2010), Santos-Rodriguez, et al (2011) and Gwang-Suk, et al (2012) that present the need for further research of this phenomenon in various geographic and industry settings. This makes the topic applicable to broad and versatile context.

In the present highly turbulent business surrounding followed by rapid Internet expansion, the customer is more knowledgeable and better informed, has easier access to larger quantum of information that provides him with more alternatives than ever which leads to wider field of opportunities for the customer in the pursuit for best value for money offers.

This situation results in easier and opportunity for more frequent change of enterprises on behalf of customers regarding the products/services they are using and to substantial harmfulness increase of the respective enterprise`s loyalty base from its competition. Earlier research statistics indicate that nearly 80 percent of company profits derive from 20 percent of its customer base. A Deloitte study published in Business Week implies that the annual churn rate of customers cost organizations in the US, \$ 15 billion a year. (Scimpaglia and Ely, 2006) Moreover, a Forbs study implies that it may cost enterprises up to 5 or 6 times more to acquire a new customer than to keep an existing one. Hence, findings from Bain and Co. indicate that 5 % increase in the company`s retention rate may result in overall profitability increase between 25 and 85 % (Reichheld and Sasser, 1990; Reichheld, F.F, 1996, Forbes, 2012) Today it is easier than ever for enterprises to lose customers and harder than ever to acquire them.

The solution to overcome the particular major challenge for the academic and real sector stands in the form of galvanization and upgrade of existing traditional marketing philosophy with new innovative conceptualization regarding Customer Relationship Management and Relationship Marketing ready to emerge in order to successfully tackle the pertinent challenges of the new business reality. It is argued that the realization of "grand" marketing theory was unattainable spurred demand for more imaginative approaches to theory building that would allow greater flexibility in establishing original marketing constructs. (Weick, 1989 and Hong and Wang, 2012) Although, Slieme, F. (2010) states that earlier research was unable to produce consensus among Academics about CRM and its benefits on behalf of businesses, emphasizing the necessity for further research in different industries and geographic settings that arise from this opinion disparity, authors Zikiene, K. (2012, "Battor and Battor (2010), Tu, et al (2011), Kabiraj and Shanmungen ((2011) Tsai-Fa and Kuo-Kuang (2009) state that building affective commitment with customers and innovation are key determinants for creation of sustainable competitive advantage for enterprises.

Enterprises that do innovate and do not position themselves in the market as customer-centric organizations will gradually lose ground to competitors. Moreover, Kabiraj and Shanmungen ((2011) and Battor and Battor (2010) state that only the creation of affective commitment between enterprises and their customers will create strong sustainable competitive advantage for the enterprise because loyal customer is not the equivalent of a satisfied customer. In this connotation, it is argued that the term "customer-centric" does not refer only to giving customers what they want but at the same time being proactive and provide customers with products and services they will want.

The airport industry is of vital importance to the countries in Southeastern Europe mainly because of geo-strategic point of view for tourism development and overall national economic growth. National citizens from these countries have extensive preferences for traveling however the weak purchasing power in the Southeastern European countries was a considerable obstacle for higher traffic rate at the respective airports.

In the past, people from the Southeast European countries have been exposed to low service quality and high ticket prices from the airlines they were flying with and airports they were flying from. Moreover, the airline competition was consisted of several airlines that use to charge extremely high prices for its services that forced people to look for other alternatives at the neighboring airports and for alternative type of transportation.

As a result, traffic rate was at an all-time low level and people mainly traveled only when they had no other option. The national airports were unable to attract low cost airlines that could make the industry more competitive and bring positive implications for national travelers in terms of lower airline traveling costs and better airline and airport service.

With the global economic crisis on the rise resulting in growing population segment with diminished purchasing power combined with people's growing interest for getaways from the ordinary everyday permanent pressure, the importance of innovation, offering world-class premium service and competitive taxes to airlines representing an essential prerequisite for offering premium service, lower ticket prices and larger destinations portfolio for passengers has become vital prerequisite for airports to position themselves as customer-centric and develop sustainable competitive advantage.

Customer-centric enterprises should focus their CRM strategies on minimization of relationships with their customers based on behavioral

commitment as it is undesired relationship by customers based on a short term notice. This commitment brings uncertainty for businesses and it is a matter of "when" and not "if", the annual churn rate (rate of annual loss of customers for enterprises) will substantially rise. In this connotation, Zikiene, K. (2012), Bulearca and Tamarjan (2010), , Tu, et al (2011), Menon and O`Connor (2007), Berry, J. (2006) and Breur, T. (2006) state that the term "behavioral commitment", referred to as "calculative" or "continuance" commitment refers to customer- marketer relationship maintained by the customer perception of high switching costs, lack of market alternatives. Academics also refer to the respective term as "customer inertia".

On the other hand, earlier research indicates that if an enterprise wants to develop a successful CRM strategy it is vital to enhance affective commitment with its customers. Moreover, authors Siems, F. (2010), Zikiene, K. (2012), Roy, S.K. (2008) and Menon and O'Connor (2007) feel that customer centric enterprises that manage to build affective commitment with their customers will be rewarded with high rate of customer retention and increase of its market share. In this connotation, Bulearca and Tamarjan (2010), Zikiene, K. (2012), Tu, et al (2011), Menon and O'Connor (2007), Berry, J. (2006) and Breur, T. (2006) agree that the term, "affective commitment" refers to a customer-marketer relationship that is built on the pillars of reciprocity and mutual trust. Furthermore, this term implies "emotional attachment" between both sides which brings them to a win-win situation based on shared values. Consequently, this research is in coherent relationship with the regional governmental strategy for tourism development and will greatly contribute to better understanding about the importance and successful paths of building affective commitment with customers and developing sustainable competitive advantage for airports. This strategy traces the path to higher airport traffic rate and larger airline competition that will positively reflect on the quality of airport and airline services that will lead to considerably lower ticket prices and improved service from the respective airlines.

The particular conceptualization will lead to higher customer satisfaction, larger inter-cultural collaboration and organizational growth that will mutually bring overall positive benefits in the form of substantial capital gains and results increase to enterprises and to all parties involved in the process. This research will be conducted by following an inductive approach using semi-structured in-depth interviews with "elite participants", comparative case studies and focus groups. In this manner, coherent integrated information and benchmarking will be presented. Considering the fact that the respective topic is unexplored emerging phenomenon, substantial space for further research is assured.

2. Literature review

CRM as a philosophy

Zikiene, K. (2012), Bulearca and Tamarjan (2010), Renart, Cabré (2008), Payne, Frow (2006), Thakur, et al (2006) and Tamošiūnienė, Jasilionienė (2007) Lawler, J.P. (2005) and Roy, S.K. (2008) agree that customer relationship management is a term that by its name itself refers to a better customer care, resulting in higher customer

loyalty and retention rate that positively impacts the results and profits of the enterprises that decide to use it as a tool. Hence, Kangal, A. (2012) and Battor and Battor (2010) feel that the proper implementation of CRM and innovation leads to creation of strong competitive advantage.

Moreover, Siems, F. (2010), Bulearca and Tamarjan (2010), Fitzgibbon and White (2005), Slemmer, M. C. (2008) and Roy, S. K. (2008) argue that CRM systems enable organizations to achieve customer focus and become more customer-centric. However, Fitzgibbon and White (2005) and Slemmer, M. C. (2008) state that in order to facilitate the management of relations with customers, software producers develop software referred to as enterprise wide applications enabling companies to view and data of any customer interaction with the particular company by detecting both side communications, recording purchases and create an understanding of each customer's preferences.

In this connotation, Slemmer, M. C. (2008), Payne and Frow (2005) and McNally, R. C. (2007) state that CRM puts companies in position to automate such processes that lead to lower sales costs and better and improved customer service collection and intelligent use of "relevant data" in order to build "superior customer experience" and "endure" relationships with customers that puts them into position to lower their annual churn rate and focus on winning new customers. Furthermore authors McNally R.C. (2007) and Rootman, C. et al (2008) feel that this process is consisted of three stages: initiation, maintenance and termination. However, authors agree that companies should attract customers in the first phase and carefully build, monitor, manage and maintain their relationships on the long term with their customers in order to achieve strong loyal relationship and avoid the third undesired phase for enterprises-termination.

Potential implementation benefits of CRM

Zikiene, K. (2012), Rao, A. (2006), Coltman, T. (2007), Kangal, A. (2012) and Battor and Battor (2010) state that advantages for organizations that build strong relations with their customers will be rewarded with maximized profits through reduced risk, better flow of information and increased number of satisfied customers through larger loyalty rate. However, Coltman, T. (2007) and Kapoulas et al (2004) feel that businesses should use CRM as a strategic tool. Coltman, T. (2007) believes that CRM should be used as "human, technical and business related activities".

On the other hand, Wachnik, B. (2012), Siems, F. (2010), Kapoulas et al (2004) and Menon and O'Connor (2007) state that enterprises should be careful when relying their strategic CRM operations solely on technological advancements because technology should be used for routine services and strive to reach personal contact as much as possible in order to put enterprises in a position to focus their operations on building strong relationships with customers as a successful path for building sustainable competitive advantage for enterprises because of the competitor's inability to imitate it in contrast to the technical innovation that is considerably easier for a competitor that is financially powerful to imitate it. Furthermore, Wachnik, B. (2012) state that his research indicates that companies are not truly aware of the benefits of CRM because they primarily use it for market survival and not as a tool for change and innovation. For this reason, authors Battor and Battor (2010) state that CRM and

innovation are widely considered to be valuable capabilities and assets in creation of sustainable competitive advantage.

However, Roy, S.K. (2008) and Lawler, J.P. (2005), Jackson, T. W. (2006) and Gessner and Volontino (2005) additionally emphasize the fact that business should not disregard the use of WEB and enable customer communication and coherent customer experience at different touch points. It is essential for enterprises to ensure coherent customer experience at different communication points between the enterprise and the customer in order to achieve standardized high quality service that the customer is used to receive one it has achieved affective commitment with the particular enterprise. Moreover, Klie, L. (2012) presents results from the Forrester Research, which indicates that 90 percent of companies lack integration across key customer interaction channels like phones, emails and especially Web self-service.

For this reason, authors Jackson, T. W. (2006) and Gessner and Volontino (2005) state that this process should be conducted by using CRM/DSS systems and business intelligence strategy (BI) that integrate different data and by effective quantitative data analysis better understand your customers that will enable enterprises to formulate their CRM strategy and customer approach more effectively. This will enable enterprises to reach higher customer retention rate by fulfilling demands at real time of customers with high quality and provide them with the opportunity to better target potential customers and make relevant offers to them that fit their demands.

However, Siems, F. (2012), Sarel and Marmorstein (2007) and Fitzgibbon and White (2005) agree that customers should be evaluated based on their lifetime value rather than be evaluated based on the metrics of an individual customer's transaction. Moreover, Berry, J. (2006) presents the following formula for calculating economic benefit of each potential communication: propensity to respond \times value of the response \div cost of communication. Furthermore, Bulearca and Tamarjan (2012), Zikine, K. (2012) and Fitzgibbon and White (2005) state that this would put organizations in position to create long-term relationships with their customers and reduce customer churn rate. In this connotation, authors Scimpaglia and Ely (2006) present facts about the importance of affective commitment for enterprises from report about Deloitte study published in Business Week implying that costs of organizations in the US that derive from the annual level of customer churn rate cost business \$15 billion.

Customer value, definition and importance of, affective, behavioral commitment

Zikiene, K. (2012) and Breur, T. (2006) emphasize that that the term "customer value" refers to the benefit or value a customer receives from using a product or service on one hand and the profit the company earns from each customer on the other. In this connotation authors Bulearca and Tamarjan (2010), Zikiene, K. (2012), Tu, et al (2011), Menon and O`Connor (2007), Berry, J. (2006), Riivari, J. (2005) and Breur, T. (2006) agree upon the fact that in order to achieve successful CRM strategy, companies should establish affective commitment with their customers and minimize the level of established relationships based on behavioral commitment with them.

Zikiene, K. (2012), Bulearca and Tamarjan (2010), Tu, et al (2011), Menon and O'Connor (2007), Berry, J. (2006), Riivari, J. (2005) and Breur, T. (2006) state that

behavioral commitment is driven by customer's perception of high switching costs whereas the desired affective commitment for enterprises focuses on perceptions of reciprocity and mutual trust for both sides that puts them into a win-win situation. Moreover, Kabiraj and Shanmungen ((2011), Battor and Battor (2010) and Tsai-Fa and Kuo-Kuang (2009) feel that only the creation of affective commitment between the organization and the customer will create strong sustainable competitive advantage for the enterprise because loyal customer is not the equivalent of a satisfied customer. However, authors Tu, et al (2011), Sarel and Marmorstein (2007) and Fitzgibbon and White (2005) agree with the previous statement but further emphasize that organizations should additionally focus on recognition to customers for their loyalty and identify the customer life cycle and customer intrinsic value to the organization and develop strong affective commitment with them. In this connotation, authors Tu, et al (2011), Battor and Battor (2010) and Scimpaglia and Ely (2006) emphasize the necessity for businesses to establish affective commitment with their customers as a vital prerequisite for building strong competitive advantage that represent successful path for business for more stable and successful future that is of great importance in a turbulent business surrounding.

Moreover, Siems, F. (2010), Zikiene, K. (2012), Bulearca and Tamarjan (2010), Scimpaglia and Ely (2006) and Tu, et al (2011) explain that business that manage to increase the level of relationships with their customers based on affective commitment on the expense of minimization of the established level of relationships based on behavioral commitment will feel benefits to its CRM strategy and business in the form of higher customer retention rate and larger customer acquisition rate that will lead to sustainable competitive advantage and considerably more positive financial results.

3. Literature gaps

1. Wachnik, B. (2012) states that considerable research is aimed towards developing models that aid Academics identify correlation between CRM and its benefits. However, the author believes that result from earlier research is not evident because of the lack of consensus among Academics and professionals regarding this argument. For this reason, the author believes that further research is necessary to uncover and determine the economic benefits of CRM implementation for the respective companies.

2. Valvi and Fragkos (2012) state that current research lacks a systematic review of the literature that provides "conceptual framework" on loyalty that would help managers understand their customers better, take advantage of industry-related factors and improve their service quality.

3. Battor and Battor (2010) state that Customer Relationship Management (CRM) and innovation are widely considered to be valuable capabilities associated with competitive advantage. However, authors feel that there is a lack of research demonstrating how they work together in order to achieve performance advantages (competitive advantage)

4. Bulearca and Tamarjan (2010) state that research findings about experiential marketing delivers overall positive experience for the customers through creation of emotional and functional value leading to long term customer satisfaction putting

companies in a more favorable market position compared to their competitors. However, authors state that more extensive research is needed in order to examine and eventually replicate these findings and arguments on different cultures and geographic regions and further investigate the connections discovered through qualitative studies.

5. Hong and Wang (2012) emphasize the theoretical necessity of reconceptualization of the constructs of trust, commitment, opportunism and relationship synergy in order to uncover the galvanizing effects of these constructs on customer loyalty. Authors encourage studies to suggest new perspectives for development of new conceptualizations rather taking previous research for granted in order to solidify earlier research findings.

6. Livne, *et al* (2011) state that the CAC (customer acquisition cost) –CR (customer retention), associated with customer loyalty (affective commitment) link is underexplored and call for research on this hot topic

7. Santos-Rodriguez, et al (2011) state that there is lack of evidence about which are the main components that are most successful for achievement of competitive advantage for enterprises and highlight the necessity for further research in this manner.

8. Gwang-Suk et al (2010) state that further research is necessary to investigate individual and contextual factors that affect consumer loyalty in a different cultural setting in a single country or multiple countries in order to achieve higher findings and data validity.

9. Siems, F. (2010) strongly encourage further research that will focus on dynamic requirement Customer Relationship Management approaches for businesses that apply life event cycle directed towards building strong sustainable loyalty with their customers

10. Bhradwaj, et al (2012) states that it is often asserted that listening to the voice of the customer more intensively (becoming more customer-centric company) can enhance the process of innovation and lead to creation of competitive advantage for the respective enterprise. However, authors emphasize the necessity of further research in this field.

11. Brown, H. (2012) state that although the effects of arrogance of relationships have been studied in the field of personality and social psychology, further research is strongly recommended to examine the effects of corporate arrogance on relationships with customers

12. Persson, A. (2011) state that the profitable management of Customer Relationship Management is currently a rapid research area. In this connotation, the author believes that further research is necessary directed onto examination of current firms practices of CRM as an asset.

13. Menon and O`Connor (2007) state that there have been calls to build both behavioral and affective commitment, but research has not yet examined how exactly this can be done.

14. Fitzgibbon and White (2005), state that there is relatively little research on attitudinal loyalty with the research mainly being focused on behavioral loyalty. They further indicate that their research was directed only to telecommunications and professional services firms. Authors recommend wider research in other services industries as well for comparison since each industry has unique features.

15. Menon and O'Connor (2007) and Fitzgibbon and White (2005) feel that research in different geographic areas is necessary as to investigate local context for any variance. They further imply that since their research is a qualitative study an additional qualitative study would be recommendable.

4. Research query

Global economic turbulence, rapid market condition variance and rising competitive power have brought enterprises to unstable grounds with rising unpredictable variables regarding future market directions and customer preferences.

The solution to overcome the particular major challenge for the academic and real sector stands in the form of galvanization and upgrade of existing traditional marketing philosophy with new innovative conceptualization regarding Customer Relationship Management and Relationship Marketing ready to emerge in order to successfully tackle the pertinent challenges of the new business reality.

The necessity for investigation of this emerging phenomenon derives from recommendations of earlier academic research by Wachnik, B. (2012), Valvi and Fragkos (2012), Hong and Wang (2012), Bulearca and Tamarjan (2012), Bhradwaj, et al (2012), Slieme, F. (2010), Brown, H. (2012), Menon and O'Connor (2007), Fitzgibbon and White (2005), Battor and Battor (2010), Santos-Rodriguez, et al (2011) and Gwang-Suk, et al (2012) that present the need for further research of this phenomenon in various geographic and industry settings. Academics highlight the necessity for organizations to focus their strategic plans and CRM strategies on continuous innovation and building affective commitment with their customers based on the premise of mutual trust and emotional attachment vis-à-vis behavioral commitment based on the concept of high switching costs.

The particular research query arises as to provide understanding this emerging phenomenon that will investigate the importance of innovation and building relationships with customers based on affective commitment in contrast to behavioral commitment and uncover their galvanizing effects on the long term customer loyalty that lead to creation of sustainable competitive advantage for Multinational enterprises owners of Airports in Southeastern Europe.

5. Aims and objectives

The respective research aims at accomplishment of the following objectives:

- 1. Understand the current established relationships between the respective enterprises and their customers. Discover the type of the respective enterprise's established relationships with its customers in the past, the type of established relationships at the moment and the how relationships could be enhanced in the future?
- 2. Uncover customers' perceptions about the level of innovation (implementation of innovative features) by the respective enterprise and the established relationship type (affective or behavioral). Furthermore, discover their attitudes about customer care and level of innovation before and after the perceptions change of the respective Airport's ownership change

- 3. Provide understanding about industry KSF-Key Success Factors that lead to creation of sustainable competitive advantage and discover if internal affective commitment among employee affects building of strong external affective commitment between the respective enterprise and its customers
- 4. Explore and provide understanding of findings for future research

6. Methodology

The choice of paradigm for researchers inevitably influences the type of findings and contributions they will arrive to. Considering the nature and the specifics of the research query and by taking into consideration past exam suggestions such as Bulearca and Tamarjan (2010), Fitzgibbon and White (2005), Menon and O'Connor (2007) and Kapoulas, et al (2002), it appears that most suitable research approach to follow is qualitative. Creswell (1994) and Kapoulas and Mitic (2012) believe that by following an interpretative research based on inductive theory- that generates "mode of inquiry" can uncover production of substantial and valuable knowledge for the business and the academic environment.

Moreover, common motives by Academic employing the interpretative approach refer to the "quest for deeper understanding of the phenomenon under investigation", "exploration of experiences" and "development of meaning" (Carson *et al*, 2001; Kapoulas *et al*, 2002; Papasolomou and Vrontis 2006; Hanson and Grimmer, 2007; Hughes *et al*, 2007; Priporas and Poimenidis, 2007; Priporas and Vangelinos, 2008 ; Andronikidis and Lambrianidou, 2010; Priporas and Kamenidou, 2011; Kapoulas and Mitic, 2012) For this reason, qualitative research is expected to produce more relevant results in marketing because it deals with descriptive and explanatory power and not with newly acquired data that are hard to find, define and often incomplete. (Gummesson, 2005, p.31)

There is mutual consent among Academics that interpretative paradigm searches for multiple representations of little acknowledged or yet to be explored or defined concepts in marketing. (Barket *et al*, 2001; Carson *et al*, 2001) Moreover, the majority of well-known marketing principles, models and theories from the early and midtwenty century were results of positivist research and aimed to establish common applicable rules and guidelines for generic strategies and marketing. (Hunt, 1983; Chung and Algaratham, 2001, Kapoulas and Mitic, 2012) Towards the end of the twentieth century there was a considerable interest in "reviving constructivism, relativism, subjectivism and interpretative approaches" in order to investigate new uninvestigated phenomena in marketing and to define new concepts and constructs. (Hunt, 1994, Hanson and Grimmer, 2007, Kapoulas and Mitic, 2012) Patience, persistence, thorough preparations supported by theoretical insights are crucial in overcoming challenges characteristic for qualitative inquiry by following interpretative approach. (Kapoulas and Mitic, 2012)

In this connotation, authors Bulearca and Tamarjan (2010) state that interpretative approach is most appropriate to follow when dealing with social phenomenon for the reason that it is not objective but highly subjective-shaped by perceptions.

By using interpretative comparative case studies the research will seek to reflect the management's understanding and perceptions about affective commitment and successful CRM strategy that leads to creation of sustainable competitive advantage for the respective enterprise. According to Hanson and Grimmer (2007) the shift towards interpretative research approach is manifested by the growing interest among researchers in using qualitative methods to uncover contemporary issues in marketing such as rising popularity of focus groups in marketing studies. The purpose of this study will be exploratory since it investigates and tries to uncover and understand an emerging phenomenon. (Kapoulas *et al*, 2002).

7. Data collection

1. Primary data will be collected from interviews that will provide detailed insight about the understanding and importance of the organization about creation of affective commitment with customers, minimization of behavioral commitment, successful implementation of CRM strategy, its meaning, potential benefits and effects that contribute to creation of sustainable competitive advantage for enterprises. In -depth interviews with "elite participants" as stated by Egan, J. (2007) is considered to be a methodology that is most likely to provide required effects in order to accomplish results in a qualitative research. Semi-structured interviews of different sessions of 45 minutes and of an hour will be conducted and will be recorded on a digital recorder for further help and documentation of data for future revisit, as it is implied in the qualitative study by (Bulearca and Tamarjan, 2010; Fitzgibbon and White, 2005). Participants in the interviews will be mainly CEOs of the enterprises and high level managers for the reason that these potential participants are considered to be highly prominent, possess the necessary expertise and well informed about the organizations' operations. (Marshall and Rossman, 1995:83).

2. In addition case studies from airports run by private investors with background screening of airport operations and results before and after the entrance of the respective management and investors will be presented for further and more explicit comparison between them. As argued by Patton (1990) that the design of a multi case uses combination of data sources in regards to secondary surveys, and company documentation to cross check data, increasing the credibility of the findings. This will provide insight about the perceptions of the respective airports` investors and management regarding the importance of building affective commitment with their customers that represent major prerequisites for implementation of a successful CRM strategy to support the strategic operations of the respective enterprise for minimization of the annual churn rate, business development and market share increase. Secondary data will be collected from various electronic sources, such as web sites, publications, relevant articles, competitors` web sites and previous research.

3. For the purpose of the research focus groups will be conducted as will feed well the findings that will be gathered to the end of the research. Focus groups will be consisted of managers and employees that the "elite participants" are comprised of, for the reason that the particular group are decision makers in the organization and possesses the necessary expertise. (Kapoulas and Mitic, 2012; Marshall and Rossman, 1995:83).

4. The method of triangulation will be applied in order to cross-check managers` and customers` perception towards behavioral and affective commitment,

suggested in the qualitative study by Kapoulas, et al (2002) This particular aim will be conducted by using documentation for triangulation. The method of documentation for triangulation will be conducted for cross-checking of relevant data. Triangulation will be done as a methodology for the reason that it is an effective path of data cross-checking collected during interviews and it increases the findings credibility. (Kapoulas and Mitic, 2012; Patton, 1990)

8. Limitations

Language

The language that will be used for the primary collection of data during the interviews will be English.

Access

Access to relevant literature on the particular topic is assured through the online bibliographic database (EBSCO, Emerald), and other relevant literature in the form of publications and books. However, primary research on the particular topic up to date has not been conducted in the concrete geographic setting that literature from other neighboring geographic regions will be accessed.

Ethical issues

For all participants, anonymity and confidentiality issues will be guaranteed through the various stages of the research process. In addition, participants will be informed about the intentions of the researcher and the research. According to personal preference, each participant will be provided with the findings of the research process. Other ethical issues do not exist that may negatively affect the research process.

Areas that will not be explored

The particular study will focus on the level of innovation and relationships that enterprise airport owners have built and would like to build with its customers in the future and how these relationships and level of innovation enables them to build customer loyalty that leads to creation of sustainable competitive advantage and will not explore political factors and their influence in this manner.

9. Conclusions

Earlier research indicates that if an enterprise wants to develop a successful CRM strategy it is vital to enhance affective commitment with its customers. Moreover, authors Siems, F. (2010), Zikiene, K. (2012), Roy, S.K. (2008) and Menon and O'Connor (2007) feel that customer centric enterprises that manage to build affective commitment with their customers will be rewarded with high rate of customer retention and increase of its market share. Considering the fact that research for the respective topic and in the particular industry has not been done yet substantial space for further research is assured.

Customer-centric enterprises should focus their CRM on minimization of relationships with their customers based on the behavioral commitment as it is a relationship that is undesired by customers and based on a short term notice. This commitment brings uncertainty for business and it is a matter of "when" and not "if", the annual churn rate (rate of annual loss of customers for enterprises) will substantially rise. In this connotation, Zikiene, K. (2012), Bulearca and Tamarjan (2010), , Tu, et al (2011), Menon and O`Connor (2007), Berry, J. (2006) and Breur, T. (2006) state that the term "behavioral commitment", referred to as "calculative" or "continuance" commitment refers to customer- marketer relationship maintained by the customer perception of high switching costs, lack of market alternatives. Academics also refer to the respective term as "customer inertia".

Enterprises that do innovate and do not position themselves on the market as customer-centric and build strong affective commitment with their customers will gradually lose ground to competitors and its market position with an end result of business termination through extensive financial losses. Moreover, Kabiraj and Shanmungen ((2011) and Battor and Battor (2010) state that only the creation of affective commitment between the organization and the customer will create strong sustainable competitive advantage for the enterprise because loyal customer is not the equivalent of a satisfied customer. In this connotation, the term "customer-centric" does not refer only to giving customers what they want but at the same time being proactive and provide customers with products and services they will want.

In this connotation, Bulearca and Tamarjan (2010), Zikiene, K. (2012), Tu, et al (2011), Menon and O`Connor (2007), Berry, J. (2006) and Breur, T. (2006) agree that the term, "affective commitment" refers to a customer-marketer relationship that is built on the pillars of reciprocity and mutual trust. In this connotation the term implies "emotional attachment" between both sides which brings them to a win-win situation based on shared values.

Academics strongly recommend enterprises to focus their business operations and CRM strategies onto creation of affective commitment as it represents pertinent prerequisite for success of the respective enterprise and creation of sustainable competitive advantage that provides them with the opportunity to become distinctive in a highly competitive market place.

This research will substantially contribute to the enrichment of academic research considering the fact that it is geographically highly relevant and it investigates a global emerging phenomenon. (Kapoulas, et al, 2002)

Furthermore, the particular research will greatly contribute to better understanding about the importance of building affective commitment with customers and developing sustainable competitive advantage for airports that will lead to higher airport traffic rate and larger airline competition that will positively reflect on the quality of airport and airline services that will lead to considerably lower ticket prices from the respective airlines.

The particular conceptualization will lead to higher customer satisfaction, larger inter-cultural collaboration and organizational growth that will mutually bring overall positive benefits in the form of substantial capital gains and results increase to enterprises and to all parties involved in the process. The particular topic represents global unexplored emerging phenomenon and for this reason, further research is necessary in various geographic and industry settings for findings validation.

Bibliography

- Andronikidis, A.I. and Lambrianidou, M. (2010), "Children's understanding of television advertising: a grounded theory approach", Psychology & Marketing, Vol. 27 No. 4, pp. 299-322
- Barker, A., Nancarrow, C. and Spackman, N. (2001), "Informed eclecticism: a research paradigm for the twenty-first century", International Journal of Market Research, Vol. 43, pp. 3-27.
- Battor, M. and Battor, M (2010) The impact of customer relationship management capability on innovation and performance advantages: testing a mediated model. *Journal of Marketing Management*. 26 (9/10) p.842-857. Full text [online]. EBSCO Industries Inc, EBSCO.
- Berry, J. (2006) The benefits of using a decision engine to optimise campaign planning for direct marketing. *Journal of Database Marketing & Customer Strategy Management.* 13 (4). p. 319-323. Full text [online]. EBSCO Industries Inc, EBSCO.
- Bhradwaj, N. Nevin, J. and Wallman, J. P. (2012) Explicating Hearing the Voice of the Customer as a Manifestation of Customer Focus and Assessing its Consequences Explicating Hearing the Voice of the Customer as a Manifestation of Customer Focus and Assessing its Consequences. *Journal of Product Innovation Management.* 29 (6). p. 1012-1030. Full text [online]. EBSCO Industries Inc, EBSCO.
- Breur, T. (2006) The importance of focus for generating customer value. *Journal of Financial Services Marketing*. 11 (1). p. 64-71. Full text [online]. EBSCO Industries Inc, EBSCO.
- Brown, H. (2012) So What If I Don't Have An iPhone? The Unintended Consequences Of Using Arrogance In Advertising. *Journal of Applied Business Research.* 28 (4). p. 555-562. Full text [online]. EBSCO Industries Inc, EBSCO.
- Bulearca, M. and Tamarjan, D. (2010) Augmented Reality: A Sustainable Marketing Tool? *Global Business & Management Research*. 2 (2/3). p. 237-252. Full text [online]. EBSCO Industries Inc, EBSCO.
- Carson, D.J., Gilmore, A., Perry, C. and Gronhaug, K. (2001), Qualitative Marketing Research, Sage, London.
- Chung, E. and Alagaratnam, S. (2001), "'Teach ten thousand stars how not to dance': a survey of alternative ontologies in marketing research", Qualitative Market Research: An International Journal, Vol. 4 No. 4, pp. 224-34.
- Coltman, T. (2007) Can superior CRM capabilities improve performance in banking. *Journal of Financial Services Marketing*. 12 (2). p.102-114. Full text [online]. EBSCO Industries Inc, EBSCO.
- Creswell, J., W. (1994) Research design: Qualitative and Quantitative Approaches, Thousand Oaks: Sage
- Egan, J. (2007) Sidelined? The future of marketing in the contemporary organization. *The Irish Journal of Management*. p.99-118. Emerald.
- Fitzgibbon, C. and White, L. (2005) The role of attitudinal loyalty in the development of customer relationship management strategy within service firms.

Journal of Financial Services Marketing. 9 (3). p.214-230. Full text [online]. EBSCO Industries Inc, EBSCO.

- Gummesson, E. (2005), "Qualitative research in marketing: road-map for a wilderness of complexity and unpredictability", European Journal of Marketing, Vol. 39 Nos 3/4, pp. 309-27.
- Gwang-Suk, Lee, G. and Park, K (2010) A Cross-National Investigation on How Ethical Consumers Build Loyalty Toward Fair Trade Brands. *Journal of Business Ethics*. 96 (4). p. 589-611. Full text [online]. EBSCO Industries Inc, EBSCO.
- Gessner, Guy H. and Volonino, L. (2005) Quick Response Improves Returns On Business Intelligence Investments. *Information Systems Management*. 22 (3). p.66-74. Full text [online]. EBSCO Industries Inc, EBSCO.
- Hanson, D. and Grimmer, M. (2007), "The mix of qualitative and quantitative research in major marketing journals 1993-2002", European Journal of Marketing, Vol. 41 Nos 1/2, pp. 58-70.
- Hong, S. and Yong, J. W. (2009) When relationship marketing collides with technology. *Journal*. 8. p. 218-230. Emerald, Group publishing limited, Emerald.
- Hughes, T., Foss, B., Stone, M. and Cheverton, P. (2007), "Degrees of separation: technological interactivity and account management", International Journal of Bank Marketing, Vol. 25 No. 5, pp. 315-35.
- Hunt, S.D. (1983), "General theories and the fundamental explananda of marketing", Journal of Marketing, Vol. 47 No. 1, pp. 9-17.
- Hunt, S.D. (1994), "On rethinking marketing: our discipline, our practice, our methods", European Journal of Marketing, Vol. 28 No. 3, pp. 13-25.
- Lawler, James P. (2005) Case Study of Customer Relationship Management (CRM) in the On-Line Affluent Financial Services Market. *Journal of Internet Commerce.* 4 (4). p.153. Full text [online]. EBSCO Industries Inc, EBSCO.
- McNally, Regina C. (2007) An exploration of call centre agents' CRM software use, customer orientation and job performance in the customer relationship maintenance phase. *Journal of Financial Services Marketing*. 12 (2). p. 169-184. Full text [online]. EBSCO Industries Inc, EBSCO.
- Marshall, C. and Rossman, G.B. (eds) (1995) *Designing Qualitative Research* (Second Edition), Thousand Oaks, CA: Sage Publications.
- Menon, K. and O'Connor, A. (2007) Building customers' affective commitment towards retail banks: The role of CRM in each 'moment of truth'. *Journal* of Financial Services Marketing. 12 (2). p.157-168. Full text [online]. EBSCO Industries Inc, EBSCO.
- Jackson, Tyrone W. (2006) CRM: From 'art to science'. Journal of Database Marketing & Customer Strategy Management. 13 (1). p.76-92. Full text [online]. EBSCO Industries Inc, EBSCO.
- Kabiraj, S. and Shanmugen, J. (2011) Development of a conceptual framework for brand loyalty: A Euro-Mediterranean perspective. *Journal of Brand Management.* 18 (4/5). p.285-299 Full text [online]. EBSCO Industries Inc, EBSCO.

- Kangal, A. (2012) Identifying the Dominant Perspective on Customer Relationship Management in the Last Decade 2000 – 2010. Journal of Business Research. 4 (4). p. 45-61. Full text [online]. EBSCO Industries Inc, EBSCO.
- Kapoulas, A, Murphy, W. and Ellis, N. (2004) The Voice of the customer in ebanking relationships. *Journal of Consumer bahaviour*. 3. p.27-51.Emerald, Group publishing limited, Emerald
- Kapulas, A. and Mitic, M. (2012) Understanding challenges of qualitative research: rhetorical issues and reality traps. *Qualitative Market Research: An International Journal.* 15 (4). p. 354-368. Full text [online]. EBSCO Industries Inc, EBSCO.
- Klie, L. (2012) Customer Service Is a Complex Ecosystem. *CRM Magazine*. 16 (1). p. 16/16. Full text [online]. EBSCO Industries Inc, EBSCO.
- Livne, G. Simpson, A. and Talmor, E. (2011) Do Customer Acquisition Cost, Retention and Usage Matter to Firm Performance and Valuation? *Journal* of Business Finance & Accounting. 38 (3/4). p.334-363. Full text [online]. EBSCO Industries Inc, EBSCO.
- Payne, A. and Frow, P. (2006) Customer Relationship Management: from Strategy to Implementation. *Journal of Marketing Management*[online], 22 (1/2),p.135-168. Full text [online]. EBSCO Industries Inc, EBSCO.
- Papasolomou, I. and Vrontis, D. (2006), "Building corporate branding through internal marketing: the case of the UK retail bank industry", Journal of Product & Brand Management, Vol. 15 No. 1, pp. 37-47.
- Payne, A. and Frow, P. (2005) A Strategic Framework for Customer Relationship Management. *Journal of Marketing*[online], 69 (4),p.167-176. Full text [online]. EBSCO Industries Inc, EBSCO.
- Persson, A. (2011) The management of customer relationships as assets in the retail banking sector. *Journal of Strategic Marketing*. 19 (1). p.105-119. Full text [online]. EBSCO Industries Inc, EBSCO.
- Priporas, C.V. and Kamenidou, I. (2011), "Perceptions of potential postgraduate Greek business students towards UK universities brand and brand reputation", Journal of Brand Management, Vol. 18 Nos 4/5, pp. 264-73.
- Priporas, C.V. and Poimenidis, I. (2008), "Services managers' awareness of crisis management:attitudes and preparation", Innovative Marketing, Vol. 4 No. 3, pp. 37-45.
- Rao, A. (2006) Multi-attribute analysis of confidence, convenience and price functions of customers of financial services firms: A GLS systems model. *Journal of Financial Services Marketing*. 10 (3). p.198-217. Full text [online]. EBSCO Industries Inc, EBSCO.
- Reichheld, F.F. (1996), The Loyalty Effect (MA: Harvard Business School Press).
- Reichheld, F.F. and W.E. Sasser, Jr. (1990), 'Zero Defections: Quality Comes to Services', Harvard Business Review, Vol. 68, No. 5 (September/October), pp. 105–11.
- Renart, L. G. and Cabré, C. (2008) Paths To Continuos Improvement Of a CRM Strategy. *Market*[online], 20 (1/2),p.61-77. Full text [online]. EBSCO Industries Inc, EBSCO.

- Rootman, C. Tait, M. and Bosch, J. (2008) Variables influencing the customer relationship management of banks. *Journal of Financial Services Marketing.* 13 (1). p.52-62. Full text [online]. EBSCO Industries Inc, EBSCO.
- Roy, Sanjit K. (2008) CRM Implementation in Banks. ICFAI Journal of Management Research. 7 (7), p.55-72. Full text [online]. EBSCO Industries Inc, EBSCO.
- Riivari, J. (2005) Mobile banking: A powerful new marketing and CRM tool for financial services companies all over Europe. *Journal of Financial Services Marketing*. 10 (1). p.11-20. Full text [online]. EBSCO Industries Inc, EBSCO.
- Santos-Rodrigues, H. Dorrego, P. and Jardon, C. M. Intellectual Capital and a Firm's Innovativeness. *Proceedings of the European Conference on Intellectual Capital*. p.378-387. Full text [online]. EBSCO Industries Inc, EBSCO.
- Sarel, D. and Marmorstein, H. (2007) Customer relationship management in banking: An introduction and strategic implications. *Journal of Financial Services Marketing*. 12 (2). p.97-101. Full text [online]. EBSCO Industries Inc, EBSCO.
- Sciglimpaglia, D. and Ely, D. (2006) Customer account relationships and e-retail banking usage. *Journal of Financial Services Marketing*. 10 (4). p.109-122. Full text [online]. EBSCO Industries Inc, EBSCO.
- Siems, F. (2010) Extending the life event cycle to relationship marketing: New implications for products and services. *International Journal of Business Research.* 10 (2). p. 81-96. Full text [online]. EBSCO Industries Inc, EBSCO.
- Slemmer, Michael C. (2008) How to Overcome Technology Obstacles. Journal of Financial Planning.p.18-19. Full text [online]. EBSCO Industries Inc, EBSCO.
- Tamošiūnienė, R. and Jasilionienė, R. (2007) Customer Relationship Management As Business Strategy Appliance: Theoretical And Practical Dimensions. *Journal of Business Economics & Management*[online], 8 (1),p.69-78. Full text [online]. EBSCO Industries Inc, EBSCO.
- Thakur, R, Summey, John H. and Balasubramanian, S. (2006) CRM As Strategy: Avoiding The Pitfall Of Tactics. *Marketing Management Journal*[online], 16 (2),p.147-154. Full text [online]. EBSCO Industries Inc, EBSCO.
- Tsai-Fa, Y. and Kuo-Kuang, C. (2009) DO MORE INVESTMENTS OR HIGHER JUSTICE WIN THE HIGHER LOYALTY? *International Journal of Electronic Business Management.* 7 (3). p.211-223. Full text [online]. EBSCO Industries Inc, EBSCO.
- Tu, Y.T. Mei-Lien, L. and Chin, H. C. (2011) The Effect of Service Quality, Customer Perceived Value and Satisfaction on Loyalty. *Journal* of Economics & Behavioral Studies. 3 (3). p.198-212. Full text [online]. EBSCO Industries Inc, EBSCO.
- Valvi, A. and Fragkos, K. (2012) Critical review of the e-loyalty literature: a purchase-centred framework. *Electronic Commerce Research*. 12 (3). p. 331-378. Full text [online]. EBSCO Industries Inc, EBSCO.

- Wachnik, B. (2012) An Analysis of the Problems Linked to Economic Evaluation of Management Support Information Systems in Poland on the Example of ERP/CRMClass Applications - Problem Analysis. Proceedings of the European Conference on Information Management & Evaluation. p.325-333. Full text [online]. EBSCO Industries Inc, EBSCO.
- Weick, K.E. (1989), "Theory construction as disciplined imagination", Academy of Management Review, Vol. 14 No. 4, pp. 516-31.
- Zikiene, K. (2012) Research of Factors Influencing Loyal Customer Switching Behaviour: Updated and Revised in the Context of Economical Crisis. *Management of Organizations: Systematic Research.* 61. p. 143-154. Full text [online]. EBSCO Industries Inc, EBSCO.

Dynamic of Social –Economic disparities and regional development in Albania.

Eva Dhimitri¹, Frederik Cucllari², Mirela Cini³

¹Phd candidate, Faculty of Economics, Department of Management, University "Fan S.Noli "Korce, Albania

evadhimitri@yahoo.co.uk

²Prof.Assoc.Dr, Faculty of Economics, Department of Finance, University "Fan S.Noli

"Korce, Albania

fcucllari@gmail.com

³Prof.Assoc.Dr, Faculty of Economics, Department of Management, University "Fan S.Noli

"Korce, Albania

mirelacini@yahoo.com

Abstract: Definition of regional disparities by connecting OECD (Organization for Economic Cooperation and Development) indicates a measure of the economic phenomenon whose intensity varies between regions within the existing state. We can study several types of disparities: economic, social, territorial. Regional disparities are distinct and extreme cases of lower economic performance and acute social problems need to be addressed in a comprehensive manner, that development assistance to regions to apply a different (or separate programs and / or by different types of levels of motivation). The paper highlight regional disparities in Albania discussing both inter-regional disparities - between four development regions corresponding to the NUTS II level and intra-regional disparities - between twelve counties(qark) corresponding NUTS III level. This paper is focused to the principles of measurement of regional disparities and used indicators in this measurement. This study focuses in four main issues: What types of disparity are addressed? What indicators can be used to measure the disparities identified? What factors underlie these disparities? Which is the role of regional policies? The aim of this paper is to show the importance of the regional policies that can ensure more balanced development of all regions, municipalities and communes.

Keywords: regional disparities, regional development policies, cohesion policy, variability rate

"Inequality is like an elephant: You can't define it, but you know it when you see it" (Fields 2001, 14)

1. Introduction

The conceptualization of regional problems varies from country to country, but there are essentially three types of disparity addressed:

• physical disparities (those associated with geographical or natural conditions);

- economic disparities (those concerned with differences in the quality or quantity of output of a region);
- social disparities (those concerned with the income or standard of living of the population).

The identification and measurement of regional disparities is fundamental to the design of policies intended to address perceived inequalities between areas. This reflects different perceptions of what constitutes a spatial problem; different approaches to selecting areas for policy intervention; different choices of indicator to identify the presence of particular characteristics; different approaches to prioritizing or combining the results of any such analysis; and different political sensitivities. Researchers suggest that disparities are usually bigger when measured on a lower level (e.g. disparities between municipalities or communes are inevitably bigger than disparities between NUTS 2 or NUTS 3 regions).

This is evident for Albania even by a quick look on regional statistics. Considerable differences in the social and economic development in Albania compound the need for a regional development policy that can ensure more balanced development of all regions, municipalities and communes. While disparities between the regions are noticeable, even greater differences in income, output, productivity and employment are observed among municipalities and communes. Regional development policy is a very new area of public policy in Albania. Formally regional development is mandated mainly under the duties of the Ministry of Economy. It is responsible for balanced regional development through harmonization of sector and regional policies. Actually there is no strategy or action plan for addressing the socioeconomic disparities and regional inequities, which would allow for more effective targeting of resources. The role of the regional or local governments in allocating and managing investments in local public services and regional economic development is extremely important. Firstly, it ensures that the priority regions are targeted, as it is much more difficult to determine the geographical use of money disbursed through the sector ministry programmed. Secondly, regional and local governments are regarded as best placed to build up the inter-municipal/governmental and public private partnerships. This implies that a major proportion of public investments should be targeted at the less developed regions, channeled through integrated regional programs and managed in partnership with regional and local governments and economic and social partners.

The experiences of new EU members show that tensions tend to arise between regional and sector approaches to investment. Therefore, the need to establish a regional development policy in Albania is manifested by a number of reasons, as confirmed by a growing understanding among various stakeholders, including the business community, the donors and central and local government institutions. There is a need to upgrade the economic infrastructure in order to create an attractive and suitable environment for future investments. A number of difficulties need to be overcome with regard to this: such as the obvious lack of resources and the legacy of the inherited administrative structures and practices by which investment planning is typically carried out by ministries with little concern for the local and regional impact. The issue of deepening regional disparities in Albania needs to be addressed better and can be done through a regional policy; and the compliance with the EU membership, which requires the regional policy and its institutions. The territory of

Albania was officially divided into 374 local government units in the first instance (65 municipalities, 309 communes and 12 regions NUTS III.). This results in a situation in which 48% of these self-governing units, representing 17% of the country's population, made up of communities with less than 5,000 residents - or, respectively, 54% and 30%, in the case of municipalities.

2. Literature review

The traditional objective of regional development policy is the reduction of territorial disparities to achieve a relative balance between economic and social levels of development in different areas of the national territory. To understand the complexity of regional development we needed a succinct presentation of economic doctrines and the influence that they had on regional development.

In the classical economic theory the efficiency advantages of the regions deriving from the comparative specialization will eventually contribute to the reduction of territorial disparities in a way that is advantageous for all the participating regions.

In the neoclassical economic theory,(The end of the 19th century and the beginning of the 20th century Herbert A. Simion; Alfred Marshall, Vilfredo Pareto) due to the presumption of the absolute mobility of the factors of production (including technology), (Peter Nijkamp; Maria Abreu "Theory of regional policies") all the inequalities in the model – embracing any kind of developmental disparities between regions – decease in the long run.

In the Keynesian economics (John Maynard Keynes; Paul Samuelson; Jean Tobin), (Albert Hirschman (1958)"The strategy of economic development") the reduction of regional disparities cannot be interpreted as the result of spontaneous market processes. The desirable processes are much more linked to the result of certain intended institutional interventions.

Endogenous growth theory interprets the productivity growth as an outcome of the spatial diffusion of knowledge ("Promoting endogenous development" Vázquez-Barquero) technology, which does not infer any automatism for the reduction of territorial inequalities. However the regional (economic) policy aiming at the deliberate development examined the connection between the concept of competitiveness and sustainability from the viewpoint of some highlighted economic theories. Endogenous factors (technology, knowledge and the internal resources of the region) can become efficient means of reducing regional disparities. (Gunnar Myrdal (1957) "Economic theory and underdeveloped regions")

New trade theory (John Friedmann (1967) "General theory of polarized development") states that the spatial variation of productivity derives from the varying levels of regional specialization, agglomeration and cluster formation. In the new institutional economics, due to the constant change deriving from the dynamic interaction of the narrowly meant economic processes and institutional conditions, the deepening or the reduction of territorial disparities can be well interpreted within the frame of the model.

The Porterian corporate strategy economics (Michael Porter: (1990)"The competitive advantage of nations")originates the regional disparities from the basic industries and clusters of the regions. Since it focuses on the "microeconomic

foundations" (the resource munificence of the region gains highlight as well), the reduction of territorial disparities characteristically does not occur through market automatisms. In an evolutionary economic (Morgan K (1997) Regional Studies) view the change in the intensity and extent of a region's innovative activities can significantly shape the regional disparities. Such changes may occur as a result of spontaneous market processes. Therefore in the evolutionary thinking the reduction of territorial inequalities through the market automatisms can be interpreted.

Each of the models presented proves its limitations, at some point, through the impossibility to identify the winning formula in view of reducing the regional discrepancies, but, each time, the other models come to complete the solutions identified within a model, by developing new dimensions and visions with impact on regional development. Each of the theories and models presented contributes to the development of integrated solutions in view of reaching the objectives set at the regional level.

3. Methodology

A key objective of regional development policy is to reduce disparities between regions and to ensure a relatively balanced level of development. To achieve this goal studies and social economic analysis based on certain techniques and methods of evaluation are necessary. In scientific literature, there are plenty of models that can be applied to assess regional disparities.

One of the methods commonly used in practice is related to the calculation and analysis of the degree of concentration/diversification of activities within a region. The territorial inequalities are the result of an uneven distribution of natural and human resources, the effect of many different economic, social, politic and demographic variables and of the way they spatially interconnect, and are also marked by the historic evolution of the regions. Consequently, the assessment of territorial disparities requires many data and information regarding various economic, social, political, cultural and geographical aspects that influence the development gaps.

The analysis of regional disparities has been based on a series of data and indicators provided by Albanian Institute of Statistics for 2000 and 2012 and Regional Offices of Statistics processed by various statistical methods. We used indicators that are standard deviation and variation coefficient. Standard deviation is the radical of a variance. Variance is the average of square deviations of the particular character value from their arithmetic averages. For adequate illustration of regional variability rate, respectively, the importance of regional differences, it is appropriate to use both variability rates, i.e. both the standard deviation and variation coefficient. Standard deviation is not a dimensionless number. It depends on the total level of the phenomenon in the country. The value of standard deviation does not express only variability rate but it can rise with the growth of the phenomenon in the country. By contrast, the variation coefficient is a dimensionless number and it shows only the value of variability (in our case of regional disparities). For that reason it is necessary to combine the two rates, especially by comparing the size of regional disparities between different regions. By comparing national differences according to different characteristics use of standard deviation does not have practical meaning and it is appropriate to use the variation coefficient.

In advanced economies regional disparities are most frequently measured by unemployment rate and the level of economic output (GDP) per capita, employment by sectors, economically active population and participation rate (economic activity rate), active and newly established enterprises (non-agriculture), foreign enterprises for 10.000/inhabitant, Credit to business 000/All. We used all of them. The applicability of the particular characteristics depends among others on the availability of quality data by the regional classification.

4. Main finding of the empirical study.

Albania is a small country in area and population it could be compared to many of the NUTS II regions in EU. Albania is lagging behind in terms of (economic) development not only from EU countries but also from most candidate and potential candidate countries. Albania is the least developed country in the region, at 26% GDP per capita of the EU27 average in purchasing power parity standard. This corresponds to 79% of GDP p.c. in Macedonia and only to 41% of GDP p.c. in Croatia. In economic structure (as reflected in the employment by sectors) Albania differs significantly from European countries. The employment in agriculture is 8 times higher than in EU 27. For individual countries the situation is similar: 3,4 higher than Croatia, 3,8 higher than Greece. (Source of data EUROSTAT, Albanian Institute of Statistic, 2012.

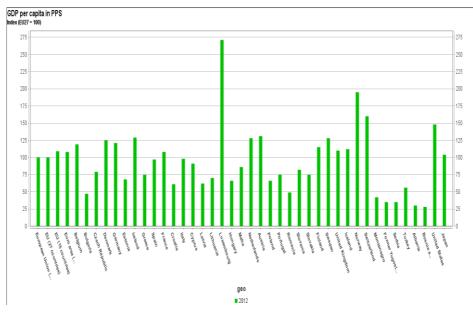


Fig 1. GDP per capita in PPS (Purchase Power Standart) EU Countries and Albania

Source: EUROSTAT, Statistics in focus 95/2012

Statistical Region Level 1	Statistical Region Level 2	Statistical Region Level 3
		Berat
		Durres
	Veri	Kukes
		Lezhe
		Shkoder
Republic of Albania	Qender	Elbasan
		Tirane
[Berat
		Fier
	Jug	Gjirokaster
		Korce
		Vlore

 Table 1. Regional structure of Albania (Division according NUTS Classification)

Based on a comparison of variation coefficient and range (min/max, expressed in % of average, number of regions significantly above and significantly below the average) lets see the results.

Indicators about demography and spatial development are: numer of population according qark, the population density, the average annual change of population, rate of urbanization and the distribution of the population accordin local government units.

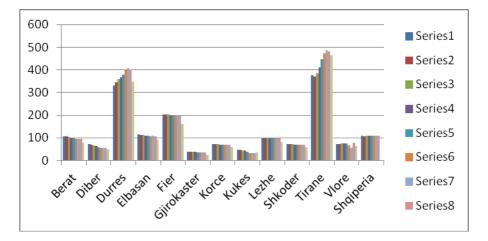


Fig.2. Dynamic of the population density 2001-2012

Increase in Tirana (+27-12,3% ne 2008,2012) Durres (+21-10,4% 2008,2012) and Vlore (+8% in 2008 and decrease 12% in 2012.); Decrease in all other qarks, especially: Kukes (-26%), Diber (-30%), dhe Berat (-29%).

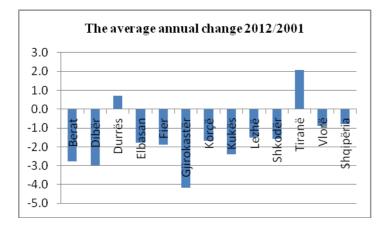


Fig.3 The average annual change in %

The rate of urbanization in country level was 42 % për vitin 2001 and 55% in 2012. Tirana had highest rate in 2001, 64 % while Dibra lowest rate with 20 %. In 2012, Durrësi was with the highest rate of urbanisation 78%, while lowest Dibra 26%.

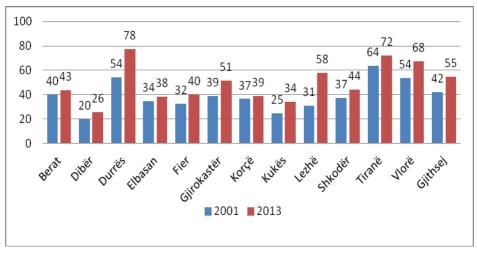


Fig 4. The rate of urbanization 2001-2013

The social economic indicators are: GDP per capita, level of employment by sector, unemployment, economically active population and participation rate, active and newly established enterprises (non-agriculture), foreign enterprises per 10.000/inhabitant, credit to business, 000 ALL/person. The level of GDP in Albania is low (10% of EU27 nominal). The economic activity as reflected by GDP is highly concentrated, only 3 regions (Tirana, Fier, Durres) generate 34% of GDP. The max/min ratio is 2,2 there are 3 regions below 75% of the average and only one is

above 125% of the average. Best performing: Tirana (112%), Durres (111%), Shkoder (109%) Worst performing: Diber, Lezhe, Kukes.(65%,74% and 78% of the average).

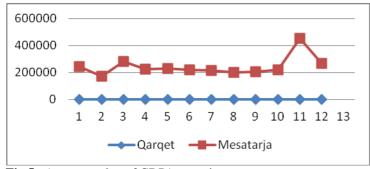


Fig.5 Average value of GDP/per capita

Level of employment by sector : We see very high employment in agriculture (44%), although declining sharply in the last years. Very high regional variation, only 2 regions could be specified as non-agricultural – Tirana (17% agricultural employment, 39% of the average) and Durres (32% of the average) – but even in these regions agricultural employment is very high in European context and 10% of employed in agriculture are in Tirana. Significantly above the average are Fier (66%), Kukes (66%), Berat (62%), Elbasan (59%), and Diber (56%). Result is decline of agricultural employment in all regions, although with different speed. The main causes are inherited economic structure, lack of other job opportunities in face of demographic changes. What are the implications? It seems not possible to sustain such high employment in view of modernization in mid-term. As already seen in the previous years the expected consequences will be pressure on the labour market, high unemployment, and migration to more urban and non-agricultural areas.

Economically active population and participation rate (economic activity rate): Overall growth of economically active population but with high oscillations (by year) and very high regional variation of growth leading to significant redistribution (higher than for the total population): migration of the active population is evident, highest growth have regions: Durres (103%), Korce (38%), Tirana (24%), Elbasan (20%), Gjirokaster (19%) highest decline have Kukes (-32%), Diber (-30%), Lezhe (-19%) The variation of participation rate in 2012 is not too high (72-113% of the average), but the extreme cases need attention: Very high in Korce (70% or 113% of the average), Very low in Diber (44% or 72% of average), Kukes (51%, 82% of the average), Lezhe (49%, 79% of average). Females are 43% of the labour force in 2012, with limited regional variation (between 40 and 48%, except in Diber -35%). However their participation rate is significantly lower than of males (53% compared to 72%) and with significant regional variation (higher than for the total participation rate).We see very high in Korce (66% or 125% of the average), very low in Diber (32% or 61% of the average), Lezhe (36% or 67% of the average), Kukes (39% or 73% of the average).

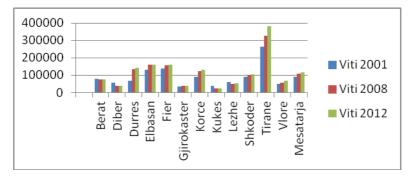


Fig 6. Distribution of active population according qarks.

Unemployment: High level of unemployment and long-term unemployment (although declining) with high and changing regional variation - The regional variation is high – between 6,1 (47% of the average) and 20,5 (157% of the average) or a max/min ratio of 3,4. The regional pattern differs in many aspects from other economic indicators: Very high unemployment - in Shkoder (20,5%), Tirane 21,8%,Elbasan 16,2%, Durres (18,2), Lezhe (17,0%) Lowest unemployment – in Diber (7,1%), Berat (6,2%), Vlore 7,01%. The reason for this pattern could be found in the dynamics of unemployment – it is declining more in regions with high unemployment levels (that could be seen as most lagging behind on other indicators) and declining less or even growing in other, more developed and urbanized regions. Most significant in Diber (from 16,3 to 7,1%), Berat (from 21 to 6,2%), Lezhe (from 29,2 to 17%). Increase in unemployment in Tirana (from 10,6 to 21,8%), Durres (from 13,4 to 18,2%) and Gjirokaster (from 12,5 to 14,4).

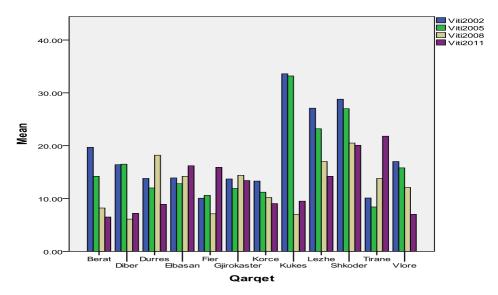


Fig 7. Level of unemployment according qarks 2002-2011

Active and newly established enterprises (non-agriculture) for 10.000 inhabitants: High dynamics of enterprises (entrepreneurship) leading to a convergence trend, but with still significant differentiation. High density in Tirana (155% of the average) as well as Durres and Vlore (122%) Lowest density in Kukes (36% of the average), Diber (39%), Lezhe (53%) and Elbasan (62%). The growth (2001-2012) was especially high in many of the low-ranking regions in 2001, e.g. Shkoder, Fier, Vlore, and Korce (2 times and more) as well as Diber and Elbasan (60-70%). The dynamics is lower for Tirana and Durres although starting from a higher base (change by 30-40%) and very low in Kukes and Lezhe. Despite the changes, however, 50% of active non-cultural enterprises are located in Tirana and Durres. The differentiation in newly established enterprises is similar but more dynamic leading to a changing and more convergent pattern. Highest density in 2012 - Tirana (144% of the average), Vlore (132%) and Shkoder (129%), slightly lower in Durres (112%) Lowest density in - Kukes (42% of average), Diber (50%), Fier (63%), Elbasan (65%) The overall growth in 2001-2010 is by 4,5 times, and all regions were growing, but especially impressive was the growth in Shkoder, Lezhe and Kukes (around 15 times), Diber (9 times), Gjirokaster (7 times).

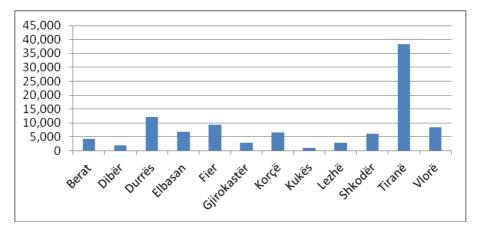


Fig 8. New active enterprises/ 10.000 inhabitants only for 2012

Foreign enterprises per 10.000/inhabitant: Extreme differentiation – by an average of 11/10.000,(one of the highest from all indicators used) Above the average are only Tirana (278%) and Durres (113%). Altogether they concentrate 80% of all foreign enterprises Only 2 regions are below but relatively close to the average – Vlore and Korce (62-63% of the average) All others are on a level between 11 and 36% of the average. Credit to business, 000 ALL/person: Extreme differentiation – higher than for foreign enterprises – Coefficient of Variation is 1,46, max/min ratio – 165. Credit to business is concentrated mainly in Tirana - value is nearly 3 times higher than the average and Tirana accounts for more than 70% of all credits. Only Durres (77%) and Vlore (55%) are closer to the average, having 11% of all credits Most of the regions are on the level of 22-33% of the average (64-96 000 ALL/person) Extreme cases – Diber (2% of the average) and Kukes (6%) as well as Berat (14%).

Overall position of the regions on economic indicators– on the basis of the set of indicators used regions could be divided in several groups Tirana – outstanding on most indicators relatively high position – Durres, Vlore Low position – Diber, Kukes, may be Lezhe All the rest – in the middle. The most disparities are related mainly to: Demography (2001-2012), population change – most expressive and clearly distinguishing the regions in 2 very different groups: 3 growing significantly (Tirana, Durres, Vlore) and 9 losing population significantly, but especially Diber (-25%) and Kukes (-29%) Population density, share of population in small Local Government Units (< 5000) more expressive than urban population, economy – not GDP but especially credits to businesses, foreign investment, Local Government Units own income, non-agriculture enterprises and newly established non-agriculture enterprises, unemployment and structure of employment (agriculture-non-agriculture), these are most differentiated as the above, but seem to be acute problems as well as cause for other problems.

Social disparities: Mainly families getting social assistance (representing the current poverty situation) (other social indicator does not show too high variation). Location and natural conditions (altitude), expressed in: Land use structure, distance/travel time to capital and to regional centre. Disparities in environmental situation and access to services are generally not so acute. More acute from them are urban waste generation (but in fact this is more related to urbanization level), cars ownership (more related to welfare) roads density – even statistically on lower is important in relation to the periphery issue.

Item	Least developed	Medium developed	Most developed
Qarks (ranked from lower to higher RD Index)	Diber Berat Kukes	Lezhe Fier Elbasan Korce Shkoder Vlore Gjirokaster Durres	Tirana
Population total (and as percentage of country's population)	391 thousands (12%)	1,998 thousands (63%)	793 thousands (25%)
Relative productivity index (%GDP to %population ratio, Albania=1.0)	0,67	0,89	1,44
Other important	Very low value	Differentiated	Highest relative

 Table 2. Classification of regions according development level.

characteristics	added in	economic	productivity
	the regional	performance, in	(about two
	economy	general	times higher than
	Significant	below the national	in the
	depopulation	average	least developed
	over the last	Stable population	regions)
	decade	over the	High population in
	(around 25-30%	last decade in the	low
	for Diber	whole	in the last decade
	and Kukes, around	group,	(about 30%
	12%	differentiated but	
	for Berat)	within +/-10%	

5. Conclusions

After analyze, we conclude that exist high divergence in economic development compared with European countries. Migration seems to be very important issue, especially in relation to high pressure on infrastructure and services in attractive areas and depopulation of some areas leading to inefficiency of development and maintenance of infrastructure and services (schools, health care, roads, water supply, etc.) that in turn will lead to diminishing attractiveness to population and business and further depopulation. Out-dated economic structure (extremely high employment in agriculture) combined with low level of urbanization. High unemployment level combined with high migration flows leading to transferring the unemployment problem in space. It seems that one of the main directions of the RD (Regional Development) programmers should be related to diversification of rural economies. Clear regional differentiation if the extremes are looked at: Most developed – Tirana and Durres, least developed (disadvantaged) - Diber, Kukes, A "grey area" in between - good on some indicators, bad on others, differences in trend etc. Competitive grant as working up to now does not address disparities, although the reformed competitive grant (Regional Development Fund) could lead to a different outcome. There is a lack of effective and efficiency of the development policies according the diversification of garks in Albania.

What's new policy aims? To promote sustainable regional development, strengthening social and economic cohesion, improve competitiveness and raising an institutional and legal framework which responds to internal development challenges and requirements for EU integration; To achieve the above objectives should be implemented both political reforms: Regionalization would mean strengthening the decentralization process, focusing on regional and local decentralization. To push forward regional decentralization, the Government should formalize the process of regionalization and to establish a clear system of regional governance in line with Autonomy Card, the Albanian Constitution and the Organic Law of local government. Level and forms of regionalization will depend on several factors that need to be considered before the decision-making. This will lead to improved services, reducing inequalities, efficiency and effectiveness of institutions, strengthening internal potentials and boost local development, employment generation and the best use of local resources.

References

- Alden, J., Boland, P., Regional Development Strategies, A. European Perspectives, London, Regional Studies Association, 1996
- Abreu M, Nijkamp P, University, Amsterdam, University of Cambridge, Cambridge, UK *Regional development policies*
- Atkinson, A. B., and A. Brandolini. (2004), "Global *World Inequality: Absolute, relative or Intermediate*?" presented at the 28th General Conference of The International Association for Research in Income and Wealth, August 22-28, Cork, Ireland.
- Fields, G. S. (2001), *Distribution and development: a new look at the developing world*, New York, London: Russell Sage Foundation; MIT Press.
- Friedmann John (1967), *General theory of polarized development* Los Angeles: University of California, The Journal of Himian Resources, Madison: University of Wisconsin.
- Gunnar Myrdal (1957) Economic theory and underdeveloped regions Gerald Duckworth & Company, Limited, 1957 - Economic development
- Hirschman Albert (1958), *The strategy of economic development*. Edition15. Publisher, Yale University Press, Original from, University of Texas.
- North Douglas (1955), *Location Theory and Regional Economic Growth* Journal of Political Economy
- Plane, D.A., L. D. Mann, K. Button, P. Nijkamp, (2007), *Classics in Planning 4: Regional Planning*, Cheltenham UK and Northampton MA USA
- Vazquez Barquero, A., "Endogenous development: networking, innovation, institutions and cities," New York, Routledge, 2002.
- Coffey, W. J., & Polese, M., "*The concept of local development: a stages model of endogenous regional growth,*" Papers of the Regional Sciences Association, Vol. 55, (1984), pp.1-12.
- Porter Michael : *The competitive advantage of nations* 1990_ Free Press, *1990* Business & Economics 855 pages.
- Wishlade Fiona and Douglas Yuill *Regional and Industrial Policy Research Paper* (1997) Number 24 Published by: European Policies Research Centre University of Strathclyde UK
- Albanian Parliament (2000). Law on Organization and Functioning of Local Governments.
- UNDP Report (2005). On Regional development in Albania; published by UNDP.
- Local Government and Decentralization Strategy Ministry of Interior; December 2006
- Local Government Services Accountability and Local Economic Growth; LGPA 2008 City Survey; USAID and IDRA
- UN Programme Annual Report, United Nation Delivering as One for Development, May 2010
- Albanian Regional Development Cross-Cutting Strategy, 2007-2013

Regional Disparities in Albania; Integrated Support for Decentralization Project "Working for Regional Development", UNDP; EU; November 2010 (http://europa.eu.int/comm/regional_policy/intro)

Albanian Statistical Institute 2008-2012 "Indicators of regions" "Regional Accounts" EUROSTAT "Indicator of European Countries" 2012

What Factors do Investments Depend on – the Case of Serbia

Dragana Draganac¹

¹PhD student Department of Business Management), University of Belgrade, Faculty of Economics, Serbia dragana c@ekof.bg.ac.rs

Abstract: Investment is a material basis for economic growth and development. Investments are prerequisites for the existence, growth and development of the corporation, as well as maintaining and improvement of competitive advantage. Profitable investments lead to an increase in the value of the corporation, contribute to the overall economic development, national welfare growth and well-being of all stakeholders. Due to its importance for the development of corporations, investment analysis was and still is a very relevant and dynamic area, which attracts a lot of attention of theorists and practitioners. This paper discusses the factors that have an effect on the amount of investment in the case of 43 Serbian PLCs (manufacturing and trading companies), that have been included in the basket of the stock index BelexLine for a time period of 3 years, 2009-2011. Regression analysis to a balanced panel is used. These are the classic panel data (short panel). The dependent variable is logarithm of investments, explanatory variables are net profit, size, depreciation and market value. This specification has adequate explanatory power. The first three explanatory variables proved to be statistically significant. The key contribution of the paper lies in the fact that we get a clearer insight into the factors that have an effect on the amount of investment in Serbia, their direction and intensity, which are the findings that may facilitate investment decision making.

Keywords: investments, balanced panel, regression analysis, Serbia.

1 Introduction

Capital budgeting decisions require investment of substantial financial resources. They are associated with a significant risk in terms of the final outcome and have long-term consequences on the corporation's operations. Today's micro and macro environment, which is characterized by growing internationalization and globalization processes, the application of new technologies and increased competition, is forcing corporations to constantly perceive changes and adapt to the new conditions by internal and external growth, as well as the divestiture of companies' parts that destroy value. An investment is any money investing that has a long term aspect, which will bring benefits in the future, and not immediately. Investing is the exchange of immediate and sure welfare, which investors give up, because they believe that they will enjoy a greater level of well-being in the future. The investment undertakings require a project approach, where a project is a planned decision for management of complex, interdependent and uncertain ventures. Evaluation and selection of investment alternatives are the most important segment of investment activity. Most companies have at their disposal a number of investment opportunities, and limited capital for this purpose, where, according to the set criteria, they choose the one alternative that will best contribute to the achievement of economic goals.

It is very important to identify the factors, i.e. explanatory variables on which the amount of investments in the corporation's fixed assets depends. For this purpose, we use regression analysis on the short panel data. This means that there is a large number of observations (corporations) observed in a shorter time period. It is necessary to determine which specification is the most adequate (pooled model, individual effects model or individual and time effects model). When the specification and evaluation model is determined, we obtain the parameters' estimate with the explanatory variables, whose interpretation and statistical significance determination are of great importance. We will determine which variables are statistically significant, which are insignificant and why, and as a result of such detailed analysis, we can conclude with greater probability whether the investment project is profitable, whether it should be accepted. These problems are much faster and easier to solve with the development of statistical software.

2 Literature Review

A large number of studies have already dealt with investments and factors that they depend on. Some of them have focused on a particular country and some on a panel of countries. Studies on the determinations of investments in the Belgium, France, Germany, and the United Kingdom were carried out by Bond, Elston, Mairesse and Mulkay (2003). Our investigation is similar to theirs. They constructed company panel data sets for manufacturing firms in Belgium, France, Germany, and the United Kingdom, covering the period 1978-1989 and investigated the role played by financial factors in each country. A robust finding was that *cash flow* and *profits* terms appear to be both statistically and quantitatively more significant in the United Kingdom than in the three continental European countries. This is consistent with the suggestion that financial constraints on investment may be relatively severe in the more market-oriented U.K. financial system.

A study carried out by Devereux and Schiantarelli (1990) investigated the relationship between investment and financial factors, like debt, stock measures of liquidity and cash flow in the environment of imperfect capital market. They grouped observations according to firm size, age, and type of industry (growing and declining). The cross-sectional variation of the impact on investment of flow and stock measures of liquidity has been analyzed also by Fazzari, Hubbard, and Petersen (1988) and by Gertler and Hubbard (1988) for U.S. firms and by Hoshi, Kashyap, and Scharfstein (1989) for Japanese firms. Gertler and Hubbard (1988) also investigate how different financial factors may affect investment. They found that the effects of capital market frictions on investment are asymmetric – having more impact in recessions than

booms. The former studies distinguish between firms according to their dividend payment behavior, while the latter classifies firms according to the strength of their institutional relationships with banks. They also concluded that particular capital market imperfections may contribute to the volatility of business output and business fixed investment, in particular. A model made by Lasky (2003) explained most of the downturn in investment between 2000 and 2003 as a result of three factors: a downturn in output relative to cyclically-adjusted total factor productivity; a rise in the real costs of funds; and below-trend growth in full-employment labor hours that left less scope for a cyclical rebound than in previous recoveries. Carlson, Fisher and Giammarino (2005) introduced to the investment problem operating leverage, reversible real options, fixed adjustment costs, and finite growth opportunities. They estimated and tested the model using simulation methods and reproduced portfolio excess returns comparable to the data.

Some other studies have focused on a large number of countries. For example, a study carried out by Love and Zicchino (2006) applied vector autoregression (VAR) to firm-level panel data from 36 countries to study the dynamic relationship between firms' financial conditions and investment. By using orthogonalized impulse-response functions they were able to separate the 'fundamental factors' (such as marginal profitability of investment) from the 'financial factors' (such as availability of internal finance) that influenced the level of investment. They found that the impact of financial factors on investment, which indicated the severity of financing constraints, was significantly larger in countries with less developed financial systems. Their finding emphasized the role of financial development in improving capital allocation and growth.

3 Data and Methodology

As we have already mentioned, in the paper we use panel data from the set of formal financial statements of 43 corporations in three consecutive years (2009, 2010 and 2011). This is a population of manufacturing and trading companies, which are included in the basket of stock index BelexLine. The dependence of investment on a set of explanatory variables is observed: *net cash flow from operating activities, net income, size, sales revenue, retained earnings, depreciation and market value*. These are short panel data. The panel is balanced, which means that there are no missing data. Pooled model, individual effects model and individual and time effects model will be estimated based on the available data.

All variables take values in the expected range. Average size of investment is 358,290,200 RSD. On average, cash flow from operating activities is 226,862,700 RSD, the average net income is 242,812,800 RSD. The average value of the companies' operating assets is 4,982,726,000 RSD. The average amount of sales is 3,746,428,000 RSD, and the average retained earnings are 158,609,900 RSD. The average annual amount of depreciation is 138,846,400 RSD, and, on average, the market value of the company is 1,940,000,000 RSD.

4 Empirical Analysis

The distribution of investment is skewed to the right, which is often the case with financial data, and has a much harder tails in comparison with the normal distribution, which can be seen in Table 1 and Figure 1.

	Percentiles	Smallest		
1%	1	1		
5%	1012	1		
10%	3133	103	Obs	129
25%	29317	107	Sum of Wgt.	129
50%	133562		Mean	358290.2
		Largest	Std. Dev.	764833.9
75%	364627	2517312		
90%	830147	3097526	Variance	5.85E+11
95%	1424562	3376723	Skewness	5.334741
99%	3376723	6581882	Kurtosis	38.42639

Table 1. Detailed descriptive analysis of the variable investment

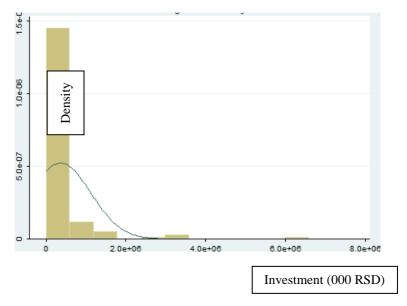


Figure 1. Distribution of the dependent variable investment

The distribution of the variable *linvestment* (logarithm of investment) is given below.

-			dependent variable)	/	
		Percentiles	Smallest		
	1%	0	0		
	5%	6.919684	0		
	10%	8.049747	4.634729	Obs	129
	25%	10.28592	4.672829	Sum of Wgt.	129
	50%	11.80232		Mean	11.22002
			Largest	Std. Dev.	2.522796
	75%	12.80663	14.7387		
	90%	13.62936	14.94611	Variance	6.364498
	95%	14.16938	15.03242	Skewness	-1.76081
	99%	15.03242	15.69983	Kurtosis	7.890963

 Table 2: Detailed descriptive analysis of the variable linvestment (logarithm of the dependent variable)

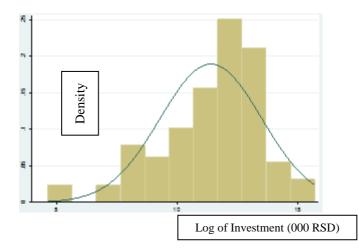


Figure 2. Distribution of the dependent variable linvestment (logarithm of the dependent variable)

The distribution of the variable *linvestment* (logarithm of investment) is skewed to the left and has a heavier tails than the normal distribution, but is much closer to the normal distribution, in comparison with the distribution of the variable *investment*. Because of this and because of the results of Box-Cox model, we opt for logarithmic transformation of the dependent variable. SK test rejects H_0 hypothesis of normal distribution of the variable *linvestment*, as well as the Kolmogorov-Smirnov test, except at the 1% level of significance.

Time-invariant variables have no variation within groups (over time), and it is a variable *id* which is used for units of observation. Individual-invariant variables have no variation between groups (between corporations) and they are *tdummy1-tdummy3*. For all other variables, there is a greater variation between corporations (between groups) than over time (within groups), so the evaluation within groups (within estimation) can cause significant losses in estimates' efficiency.

Correlation matrix of variables that includes a significance level of correlations shows that, at the 5% significance level, all explanatory variables are statistically significantly positively correlated with the *logarithm of investment*. The dependent variable *logarithm of investment* is the most correlated with the explanatory variable *revenue*, and then with *depreciation*, *size1* (size), *nd* (net profit), *market value*, *retained earnings*. It is the least correlated with the dependent variables that are negatively correlated with the dependent variable.

Firstly, we will estimate pooled model and get pooled OLS estimation, which has the incorrect standard errors because of the possible presence of heteroscedasticity and autocorrelation. However, 3 out of 7 explanatory variables are significant at the 10% significance level: net profit, depreciation and market value. Therefore we need to repair the model specification, and we choose a model with four explanatory variables: net income, size, depreciation and market value.

Source	SS	df MS		Number of o	bs = 129	
					F(4, 124) =	14.47
Model	259.214795	4	64.8036988		Prob > F = 0	0.0000
Residual	555.440919	124	4.47936225		R-squared =	0.3182
					Adj R-squar	ed=0.2962
Total	814.655714	128	6.36449777		Root MSE	= 2.1165
linvestment	Coef.	Std. Err.	t	P > t	[95% Co	nf.Interval]
nd	1.62E-06*1	6.79E-07	2.38	0.019	2.75E-07	2.96E-06
size1	1.55E-07*	6.35E-08	2.43	0.016	2.89E-08	2.80E-07
depreciation	4.90E-06*	1.83E-06	2.68 0.008		1.28E-06	8.51E-06
market_value	-1.61E-10*	9.86E-11	-1.64	0.104	-3.56E-10	3.39E-11
_cons	9.68978	0.2861772	33.86	0	9.123355	10.25621

Table 3: Pooled OLS model with 4 statistically significant explanatory variables

¹ Legend: We use asterisks (*) to show the statistical significance of the variables.

It is seen that this is better model specification. The model is statistically significant: (F (4,124) = 14:47, p value = 0.0000. Value of determination coefficient is 0.3182 and adjusted determination coefficient is 0.2962. However, it can be said that all explanatory variables are significant at the 10% significance level.

We will test the existence of heteroscedasticity because there is the large number of observations (n = 43). H_0 hypothesis about homoscedasticity of random errors is rejected at the 5% significance level by testing heteroscedasticity using squared residuals vs. estimated log investment However, we cannot reject H_0 at the 1% significance level (p-value is 0.0332). Also, we cannot reject H_0 about homoscedasticity of random errors both at the 1% and 5% significance level (p-value is 0.3249) by testing heteroscedasticity using squared residuals vs. explanatory variables. Thus, we conclude that the random errors are homoscedastic.

Further we will test the existence of residuals' autocorrelation. Residuals are serially correlated (grouped by corporations). Hence, it is necessary to use cluster-robust standard errors.

We will determine the within groups estimator, which is consistent in the FE and RE model, except that other estimators are more efficiently in the RE model. FE estimation given below eliminates individual effects and individual variables during the estimation process.

The obtained results can be summarized as follows. First, the individual effects are eliminated, but we have information on the average of individual effects (cons = 11.57). There are no individual variables, and there are no problems related to the estimation of their coefficients. Second, none of the regresors have statistically significant effect on investment both at 5% and 10% level of significance. Third, because of the only reliance on the variation within groups in the estimation process, obtained standard errors are, as we expected, mostly larger than the standard errors obtained by Pooled OLS estimation, which is based both upon the variations within groups and the variations between groups. Fourth, individual effects are statistically significant, at the 5% significance level which is confirmed by conducted F test: (F (42,82) = 5.14), p value = 0.0000, H₀ is rejected. Fifth, the regressors are, to some extent, correlated with the individual effects (corr (u_i , Xb) = -0.56), which can lead to inconsistency of the Pooled estimation, between groups estimation and RE estimation. In these circumstances, FE estimation is consistent because it eliminates the individual effects and the problem of (single) endogenousity of regressors. Thus, a positive feature of FE estimation is its consistency (even in conditions of correlation of regressors with individual effects), while its negative features are already mentioned: loss of efficiency and the inability to estimate both the effects of individual explanatory variables (which are not present in the model) and the individual effects on linvestment. The latter deficiency can be eliminated by using LSDV (least-squares dummy variables) estimation within a group that does not eliminate the individual effects and individual variables. Results of WITHIN and LSDV estimation coincide. The whole regression is significant, but none of the regressors. The FE estimation with cluster-robust standard errors are assessed. We will use cluster-robust standard errors. They are less than previously calculated standard errors.

Estimating the FE AR (1) model, we conclude that there is no autocorrelation under the AR (1) scheme, because the value of BFN-DW statistic equals to 2.0285916. F test says that the individual effects are significant. (F (42,39) = 6:30, p value = 0.0000).

Now we will define the estimation between groups, which is consistent when the RE model is adequate and inconsistent when the FE model is adequate. As we can see, in our case estimation between groups is much closer to the Pooled OLS estimaton than to the FE estimaton within groups.

Between regression	(regression on group		Number of o	bs =	129	
Group variable: id				Number of g	roups =	43
R-sq:	within = 0.0011					
	between = 0.4258			Obs per grou	p: min =	3
	overall = 0.3058				avg =	3.0
					max =	3
				F(4,38)	=	7.04
sd(u_i + avg(e_i.))=	1.824633			Prob > F	=	0.0002
linvestment	Coef.	Std. Err.	t	P > t	[95% Conf	.Interval]
nd	2.85E-06*	1.37E-06	2.07	0.045	6.64E-08	5.63E-06
size1	1.59E-07	9.83E-08	1.62	0.113	-3.95E-08	3.58E-07
depreciation	4.55E-06	2.79E-06	1.63	0.112	-1.11E-06	0.0000102
market_value	-2.62E-10	1.66E-10	-1.58	0.123	-5.99E-10	7.47E-11
_cons	9.610921	0.4318976	22.25	0.000	8.736591	10.48525

Table 4: Between estimator

Let's turn to the random effects model (RE). RE estimation is consistent if the RE model is adequate and inconsistent if the FE model is adequate. FGLS estimation of the RE model will be determined, i.e. FREGLS estimation, which is based both on variations within groups and variations between groups.

Random-effect	Random-effects GLS regression				Number of obs $= 129$		
Group variable	: id			Obs per group: $min = 43$			
R-sq: within =	R-sq: within $= 0.0001$				oup: $\min = 3$		
between =	= 0.3848				avg = 3.0		
overall = 0.3060				max = 3			
Random effects	Random effects u_i ~ Gaussian						
corr(u_i, X)	$corr(u_i, X) = 0$ (assumed)						
theta	= .567806	56					
linvestment	coef.	Std. Err.	z	P> z	[95% Conf.]	[nterval]	
nd	6.50E-07	5.91E-07	1.1	0.272	-5.09E-07	1.81E-06	
sizel	1.23E-07	8.51E-08	1.45	0.148	-4.37E-08	2.90E-07	
depreciation	5.04E-06*	2.55E-06	1.98	0.048	3.86E-08	0.00001	
market_value	-5.02E-11	1.15E-10	-0.44	0.662	-2.75E-10	1.75E-10	
_cons	9.846356	0.4181386	23.55	0.00	9.0268190	10.66589	
sigma_u	1.6454196						
sigma_e	1.3658859						
rho	0.592035	(fraction of v	ariace du	e to u_i)			

Table 5: FREGLS estimator

After the implementation of Breusch-Pagan's individual effects test and Honda individual effects test, we can summarize the results as follows. First, the individual effects are statistically significant at the 5% significance level (chi2 (1) = 38.65, p value = 0.0000), which is also confirmed by BP test (chi2 (1) = 32.25, p = 0.0000) and Honda test (5.6789083, p = 0.0000). Second, consistency of the FREGLS estimation is based on the assumption of the absence of correlation between regressors and individual effects, which is not fulfilled in our case. Third, theta coefficient equals to 0.5678 and indicates that FREGLS estimation is, to some extent, closer to FE estimation within groups than Pooled OLS estimation. Fourth, FREGLS estimation assumes constant autocorrelation coefficient of composite errors, regardless of the order of delay, which equals to 0.59 (rho) according to the obtained results. Neither this assumption is fulfilled in our case, as we have already noted that there is autocorrelation which decreases with an increase in the number of delays. With the aim to relax the assumption we use cluster-robust standard errors.

Random-effects GLS regression Number of obs = 129						
Group variable	: id			Obs per group: $min = 43$		
R-sq: within =	= 0.0001			Obs per	group: $min = 3$	
between =	= 0.3848				avg = 3	
overall =	= 0.3060				max = 3	
Random effects u_i ~ Gaussian						
corr(u_i, X)	=0 (assu	med)				
theta	= .56780	656				
(Std. Err. adjus	ted for 43 clust	ers in id)				
		Robust				
linvestment	coef.	Std. Err.	Z	P> z	[95% Conf. I	nterval]
nd	6.50E-07*	3.92E-07	1.66	0.097	-1.19E-07	1.42E-06
sizel	1.23E-07*	6.14E-08	2.01	0.045	2.84E-09	2.43E-07
depreciation	5.04E-06*	1.58E-06	3.19	0.001	1.95E-06	8.13E-06
market_value	-5.02E-11	1.00E-10	-0.5	0.616	-2.46E-10	1.46E-10
_cons	9.846356	0.5151346	19.11	0.00	8.836711	10.856
sigma_u	1.6454196					
Sigma_e	1.3658859					
rho	0.592035	(fraction of	variance due	to u_i)		

Table 6: FREGLS estimator with cluster-robust standard errors

FREGLS estimation has less standard errors than the FE estimation within groups, as is expected given the fact that, unlike the FE estimation within groups, it is based both on the variations in the group and the variations between groups. At the 10% level of significance, significant explanatory variables are as follows: *net income, size* and *depreciation*. FREGLS evaluation of the RE model is below followed by ML (Maximum Likelihood) estimation.

We notice the following. First, the standard errors have been significantly increased, which is surprising, and are usually larger than standard errors of within FE estimation. Second, we can reject H₀ about the absence of individual effects based on the LR test at the 5% significance level (chibar2 (01) = 39.26, p value = 0.000).

It will be determined which specification is better, FE or RE, i.e. whether there are fixed or random individual effects by using the Hausman's test.

	Coefficients					
	(b)	(B)	(b - B)	sqrt (diag (V_b -		
	WITHIN	FREGLS	Difference	B_B)) S. E.		
	3.06E-					
nd	07	6.50E-07	-3.44E-07	2.92E-07		
	-8.26E-					
sizel	08	1.23E-07	-2.06E-07	1.80E-07		
	-3.40E-					
depreciation	08	5.04E-07	-5.07E-06	7.45E-06		
	-2.56E-	-5.02E-				
market_value	12	11	4.77E-11	1.45E-10		
$b = \text{consistent under Ho and Ha; obtained from xtreg}$ $B = \text{inconsistent under Ha, efficient under Ho; obtained from xtreg}$ $Test: Ho: difference in coefficients not systematic$ $chi2(3) = (b-B)'[(V_b-V_B)^{(-1)}](b-B)$ $= 5.71$ $Prob>chi2 = 0.1267$						

Table 7: Hausman's test; FE vs. RE specification

 H_0 cannot be rejected at the 5% level of significance, which means that the differences between FE and REGLS estimations are not statistically significant (chi2 (3) = 5.71, p = 0.1267). The assumption that there is no correlation between repressors and individual effects is fulfilled. FE estimations include only the differences within the groups, and RE estimations both within and between groups. We choose the random effects model, because REGLS estimations are more efficient than FE estimations. We will show the robust Hausman's test, which is used when the assumptions of the standard Hausman's test are not fulfilled.

Table 8: Robust Hausman's test

	Observed Coef.	Bootstrap Std. Err.	Z	P> z	1.01111	l-based f.Interval]
nd	-3.44E-07	4.03E-07	-0.85	0.393	-1.13E-06	4.46E-07
sizel	-2.06E-07	1.93E-07	-1.07	0.286	-5.83E-07	1.72E-07
depreciation	-5.07E-06	6.33E-06	-0.80	0.423	-0.0000175	7.34E-06
market_value	4.77E-11	1.35E-10	0.35	0.723	-2.16E-10	3.12E-10
_cons	1.720492	1.186077	1.45	0.147	-0.6041755	4.04516
chi2(4) = 4.26						
Prob > chi2	2 = 0.3726					

The application of robust Hausman's test gives the same conclusions. H_0 , cannot be rejected at the 5% significance level which means that the differences between FE and REGLS estimations are not statistically significant (chi2 (4) = 4.26, p = 0.3726). We choose the random effects model, because REGLS estimations are more efficient than FE estimations.

The presence of time effects will be tested by application of the F test of time effects in the FE model and the BP test of time effects in the RE model, although it is less likely that there are time effects. F test statistic is 0.29873042, p value is 0.74230, and it can be concluded that we cannot reject H₀ about no presence of time effects at the 5% significance level. BP test statistic is 1.54, p value is 0.2153, and the conclusion is the same as previous.

Let's summarize the results of the empirical analysis. Initially 7 explanatory variables were included in the model but after various specification searches the model that best explains *investment* is FREGLS model. Significant explanatory variables are: *net income, size* and *depreciation*.

5 Conclusions

The main objective of the paper was to determine the factors on which the amount of investments in Serbia depends and how these findings affect the investment decision making process. We concluded that statistically significant explanatory variables are net income, depreciation and size. A large number of variables are statistically insignificant, which is surprising. Investment funds are primarily collected from internally generated cash flow from operating activities, and this variable proved to be statistically insignificant. Also, from the accounting point of view, retained earnings, as the part of net income left over after payment of dividends to shareholders, represent the largest and most important source of companies' internal financing. However, in our model, both *net income* and *retained earnings* are statistically insignificant variables. The market value of the company proved to be statistically insignificant explanatory variable. Bearing in mind that the units of observation are enterprises that are part of BelexLine stock index, which is a general index of the Belgrade Stock Exchange, it is not surprising that this variable proved to be statistically insignificant, given that a large number of shares are not traded at all, some are traded with only rarely and only a few shares are active. Therefore, the explanation is underdevelopment of domestic capital market.

F test, BP test and Honda test show that the individual effects are statistically significant. Individual effects model is evaluated in fixed and stochastic specification. A Hausman's test showed that there is no significant difference between the FE estimation within groups (FE model) and FREGLS estimation (RE model), from which it was concluded that the repressors are not correlated with the individual effects.

Since the RE estimation includes both variations within and between groups, and the FE estimation includes only variations within groups, the conclusion is that the RE estimation is more effective in comparison with the FE estimation, so we opted for the RE specification and FREGLS estimation. The FE and RE (FREGLS) estimations with cluster-robust standard errors are given. F test and BP test show that the time effects are not statistically significant. Finally, it can be said that the RE specification and FREGLS estimation are appropriate, the dependent variable is *ln of investment*, explanatory variables are *net profit* (nd), *size* (size 1), *depreciation* and market value. This specification is adequate because the overall determination coefficient is 0.3060. The first three explanatory variables are statistically significant.

The question is how many corporations that operate in Serbia conduct such analyses in investment decision making process. There is a problem of information asymmetry and, as a consequence, an overinvestment problem, where managers are expanding their 'empires', investing in unprofitable investment projects and destroying free cash flows. It is necessary to improve the quality of corporate governance, develop the capital market, improve the investment environment, and the like. Our data are unreliable, time series is very short (in our case, only 3 years' time series) and this is a major problem in determining which factors investments depend on. Also, the data we used in empirical research have been obtained from the official set of financial statements and have all shortcomings that are inherent to accounting data. Only the data of PLCs that are included in a stock market exchange index are used and other legal forms of businesses are not considered. 43 units of observation are both theoretically and practically a small number for the purpose of panel analysis. If the data are unreliable, the research findings and decisions validity are questionable.

It is necessary to improve studies and empirical researches about factors on which investments depend. Also, it is of the utmost importance to increase the requirements for information disclosures necessary for the analysis. Longer data time series should be created. Serbian experts and practitioners should monitor trends in investment analysis of developed countries. It is very important to examine the relationship between the level of investment, on the one hand, and debt, liquidity, solvency and profitability indicators, corporation market value and dividend policy, on the other. This should be done on a regular basis and will help making wise investment decisions that will increase the value of the corporation.

References

- Agarwal, R.N. (1997) Foreign Portfolio Investment in Some Developing Countries A Study of Determinants and Macroeconomic Impact, *Indian Economic Review*, XXXII, (2), p. 217-229
- Baker, H.K., English, P. (2011) Capital Budgeting Valuation: Financial Analysis for Today's Investment Projects (Robert W. Kolb Series), John Wiley&Sons, Inc., Hoboken, New Yersey
- Bond, S.R., Elston, J.A., Mairesse, J., Mulkay, B. (2003) Financial Factors and Investment in Belgium, France, Germany, and the United Kingdom: A Comparison Using Company Panel Data, *Review of Economics and Statistics*, 85, p. 153-165
- Cameron, A.C., Trivedi, P.K. (2010) *Microeconometrics Using Stata*, Stata Press, Lakeway Drive, College Station, Texas

- Carlson, M., Fisher, A., Giammarino, R. (2005) Corporate Investment and Asset Price Dynamics: Implications for the Cross-section of Returns, *The Journal of Finance*, 59, (6), p. 2577–2603
- Dayadana, D., Irons, R., Harrison, S., Herborn, J., Rowland, P. (2008) *Capital Budgeting: Financial Appraisal of Investment Projects*, Cambridge University Press, New York
- Devereux, M., Schiantarelli, F. (1990) Investment, Financial Factors, and Cash Flow: Evidence from U.K. Panel Data, in Hubbard, R.G. (eds), Asymmetric Information, Corporate Finance, and Investment, University of Chicago Press, Chicago, p. 279 – 306
- Environmental, Social, and Governance Factors at Listed Companies: A Manual for Investors, CFA Codes, Standards, and Position Papers, May 2008, (2)
- Fazzari, S.M., Hubbard, R.G., Petersen, B.C. (1988) Financing Constraints and Corporate Investment, NBER Working Paper No. 2387, National Bureau Of Economic Research, Cambridge, p. 9-16
- Gertler, M., Hubbard, R.G., Kashyap, A. (1991) Interest Rate Spreads, Credit Constraints, and Investment Fluctuations: An Empirical Investigation, NBER Chapters, in: *Financial Markets and Financial Crises*, National Bureau of Economic Research, Inc., p. 11-32
- Gertler, M., Hubbard, R.G., Kashyap, A. (1988) Financial Factors in Business Fluctuations, National Bureau of Economic Research, Inc., p. 12-16
- Hoshi, T., Kashyap, A., Scharfstein, D. (1989) Bank Monitoring and Investment: Evidence from the Changing Structure of Japanese Corporate Banking Relationships, in Hubbard, R.G. (eds), Asymmetric Information, Corporate Finance, and Investment, University of Chicago Press, Chicago, p. 110-116
- Hoshi, T., Kashyap, A., Scharfstein, D. (1991) Corporate Structure, Liquidity and Investment: Evidence From Japanese Industrial Groups, *Quarterly Journal of Economics*, 106 (1), p. 33-60
- Ivanisevic, M. (2012) *Poslovne finansije*, Publishing Center of the Faculty of Economics, Belgrade
- Lasky, M. (2003) A Putty-clay Model of Business Fixed Investment, Congressional Budget Office, Washington, DC
- Love, I., Zicchino, L. (2006), Financial development and dynamic investment behavior: Evidence from panel VAR, *The Quarterly Review of Economics* and Finance, 46, p. 190–210
- Shapiro, A.C. (2005) *Capital Budgeting and Investment Analysis*, Pearson Prentice Hall, Upper Saddle River, New Yersey

Proactive Feedback-Seeking Behaviour as a Valuable Resource for Nascent Social Entrepreneurs: Entrepreneurial Opportunity Evaluation and Refinement

Andreana Drencheva1

¹ Institute of Work Psychology, Management School, University of Sheffield a.drencheva@sheffield.ac.uk

Abstract. This article reviews the process of social entrepreneurship as a hybrid phenomenon that poses many challenges to the design of social ventures and the sustainability of the change they create. These challenges are most salient in the nascent stages of the process when social opportunities, expressed as business ideas, are evaluated, refined, and just in the beginning of the exploitation stage. Designing a business model that facilitates creating and sustaining an operational venture and social change in mutually reinforcing ways is a complex endeavour that isn't the result of a single insight, but product of experimentation, adaptation and refinement. Proactive feedback-seeking behaviour is proposed as a valuable entrepreneurial resource that could be used to address the challenges in the early stages of the process by decreasing uncertainty, minimising risk, setting goals, and increasing innovation, social capital and access to resources. The concluding discussion focuses on the proposed methodology, its limitations, and the theoretical and practical contributions of this research project as facilitating both the social and the entrepreneurial in social entrepreneurship.

Keywords: social entrepreneurship, feedback-seeking behaviour, entrepreneurial opportunities, nascent entrepreneurship, entrepreneurial resource

1. Introduction

Socially-minded and innovative individuals like Florence Nightingale (PBS, 2005), Robert Owen (Shaw & Carter, 2007), and Muhammad Yunus (Yunus, 2007), among many others, have been active members of their local and global communities for a long time. However, this phenomenon of addressing social issues with innovative business ideas, labelled as social entrepreneurship, has been gaining more attention from the media, policy-makers, investors, and academics only in the past three decades. In this short period, a number of conferences and streams within established conferences together with various university courses and programmes have emerged, a wide range of books and articles have been published, and a variety of support organisations such as the Skoll Foundation and the Acumen Fund have been created. Even policy-makers have recognised the importance of social entrepreneurship and have become actively engaged by creating various legal forms (e.g. in the UK, USA, Italy, South Korea, France, Belgium, Portugal) and support services (e.g. Office of Social Innovation and Civic Participation in the United States and the European Commission's Social Business Initiative) to meet the needs of social enterprises.

Albeit these efforts to establish a hospitable environment, creating a new organisation is always a challenging activity, especially for hybrid organisations such as social enterprises. By their nature, socially entrepreneurial organisations pose many challenges to the design of business models that facilitate the creation of sustainable ventures and social transformation. Because of the characteristics of the process, nascent social entrepreneurs operate in an environment characterised by high uncertainty, ambiguity, and complexity, thus the creation of a viable social business model is not the result of a single insight of a social opportunity, but the product of continuous learning, adaptation, and refinement. In this environment, proactive feedback-seeking behaviour is a valuable entrepreneurial resource that could be used to address the challenges of nascent social entrepreneurs.

The next section provides a review of the social entrepreneurship literature and a definition. The author discusses the challenges that nascent social entrepreneurs face and proposes the concept of proactive feedback seeking as an entrepreneurial resource that could be used to address these challenges. The last sections of the paper offer a short discussion of the proposed methodology and concludes with the limitations and contributions of the research.

2. Literature review

Social entrepreneurship

Despite its growing recognition, social entrepreneurship remains a complex topic with unclear conceptual boundaries. It means different things to different people (Dees, 1998) because of the breadth of manifestations of the phenomenon created by social bricoleurs, social constructionists, and social engineers (Zahra et al., 2009) across industries, sectors, and geographies.

One aspect that remains consistent across different manifestations and definitions of the phenomenon is the commitment and priority given to a mission of creating social wealth through the delivery of products or services in a sustainable manner (Lepoutre et al., 2013; Mair & Martí, 2006; Mair & Schoen, 2007; Seelos & Mair, 2005; Smallbone et al., 2001; Tracey & Phillips, 2007). This characteristic has three interdependent aspects: priority given to social wealth creation, direct action toward creating social change, and sustainability of the organisation.

Unlike other social wealth creators (e.g. social activists) that stimulate social change through rhetorical and persuasion strategies that influence others to act, social entrepreneurs create social change by acting to develop products and services that directly address the societal problem (Bacq & Janssen, 2011). Creating and maintain-

ing a sustainable venture, alliance or initiative that relies on self-sufficiency through economically sustainable solutions is essential for the development and delivery of such products and services because it is a more reliable approach to creating social change than the traditional charity model (Mair & Schoen, 2007; Tracey & Phillips, 2007; Weerawardana & Sullivan Mort, 2006). This emphasis on creating and distributing products and services for social good puts social entrepreneurs between social activists and commercial entrepreneurs. However, profit is the enabler for achieving the social mission (Wilson & Post, 2013), thus, the difference between purely commercial and social ventures lies in the priority given to economic value accumulation versus social value creation, which might include economic value (Santos, 2012).

As social ventures often address complex societal issues, such as poverty, unemployment, lack of access to education or healthcare, in resource-scarce environments (Austin, Stevenson & Wei-Skillern, 2006; Lumpkin et al., 2013), innovation is not just a mere requirement to consider the venture entrepreneurial (Schumpeter, 1961), but an essential characteristic that allows for the venture to be social. In broad terms, innovation in social enterprises addresses primarily two strategic issues: capital raising and delivery of services (Weerawardena & Sullivan Mort, 2012). Innovation can result in the invention of new products, in constantly finding new ways to structure, fund and market programmes, in accessing, combining and utilising resources, and in creating new business models (Alvord, Brown & Letts, 2004; Dees, 1998; Lepoutre, et al., 2013; Sullivan Mort, 2012). Thus, innovation is at the core of social enterprises, not only to create sustainable organisations that survive in turbulent environments, but also to effectively and sustainably address the complex social issues they attempt to solve.

However, innovation can also be detrimental to the emergence of the venture as it requires educating customers, building legitimacy, and larger resource commitments (Koellinger, 2008; Samuelsson & Davidsson, 2009; Shepherd, Douglas & Shanley, 2000). An analysis of PSED II data (Renko, 2012) showcases this complexity. Although nascent social entrepreneurs consider their ideas to be more novel compared to the competition than nascent commercial entrepreneurs do, they struggle to create operational ventures. This association is even stronger when the social enterprise's offering is new to the market (Renko, 2012).

Just like commercial entrepreneurs, social entrepreneurs create innovative business models, products, services, initiatives, and strategies that can be employed in both emerging and existing organisations. Social entrepreneurship is a process (Mair & Martí, 2006) of applying real-world problem-solving creativity (Bornstein, 1998) in recognising, evaluating, shaping, exploiting, and scaling opportunities (Dees, 1998; Mair & Martí, 2006; Mair & Noboa, 2006; Peredo & McLean, 2006; Perrini, Vurro & Constanzo, 2010; Shaw & Carter, 2007; Thomspon, Alvy & Lees, 2000). It is a process that relies on consistency between individual, organisational and contextual elements (Perrini, et al., 2010) because contextual factors such as local regulatory, normative and cognitive mechanisms shape socially entrepreneurial opportunities, venture emergence, and strategic orientation (Seelos, Mair, Battilana, & Dacin, 2011). Viable socially entrepreneurial opportunities are created by the convergence of the ability to access or develop economic, environmental and social resources, a proposition to introduce new goods, services, raw materials, markets, and/or means-ends re-

lations that generate value for the community, and sustained individual or collective motivation to pursue the opportunity (Lumpkin et al., 2013; Murphy & Coombes, 2009; Seelos, et al., 2011).

By their nature, socially entrepreneurial opportunities usually exist in resourcescarce environments (Austin et al., 2006; Lumpkin et al., 2013). However, access to the appropriate political, human, financial or social capital (Mair, Battilana & Cardenas, 2012) is a main constraint for generating social wealth (Weerawardena & Sullivan Mort, 2006) because organisational sustainability depends on balanced allocation of resources toward business action that builds and maintains competitive advantage and toward social action that builds and maintains organisational legitimacy (Moizer & Tracey, 2010). Social entrepreneurs leverage different strategies to address resource needs, such as using socially-embedded resources (Stryjan, 2006), combining resources (Mair & Martí, 2009; Mair & Noboa, 2006) in new ways (Mair & Martí, 2006), mobilising and building resources (Alvord, et al., 2004; Seelos & Mair, 2007; Thompson et al., 2000) and forming cross-sector partnerships (Sakarya et al., 2012), which creates a multi-agency environment (Shaw & Carter, 2007).

While resources are allocated toward building and maintaining legitimacy, legitimacy itself facilitates access to resources and proves social value (Suchman, 1995; VanSandt, Sud, & Marme, 2009). As social ventures differ in legal forms and blend values, structures and practices from different institutional domains to produce new organisational forms, they struggle to gain cultural alignment, perform equally well according to all institutional logics, and establish legitimacy across multiple institutional domains (Battilana & Dorado, 2010; Galaskiewicz & Barringer, 2012). They may be perceived as too commercial by actors in the non-profit sector and not commercial enough by actors in the for-profit sector (Galaskiewicz & Barringer, 2012). In addition, hybrid organisations that accommodate conflicting institutional logics can face issues related to organisational identity. Tensions and conflicts between logics can make the organisation unstable and unlikely to retain its hybrid identity (Battilana & Dorado, 2010).

Social enterprises, just like any organisation, aren't passive in the complex process of acquiring resources, building stable organisational identities, and gaining legitimacy. Instead they are actively engaged in a negotiation and renegotiation of social orders and institutional logics through bricolage, cross-sector partnerships, communication channels, such as speaking, writing and theatre performance, and rhetoric strategies, such as drawing on socially accepted meta-narratives, translating, and using analogies and metaphors (Desa, 2012; Di Domenico, Haugh & Tracey, 2010; Mair & Martí, 2009; Mair, Martí & Ventresca, 2012; Sakarya et al., 2012).

Beyond gaining and negotiating legitimacy, social entrepreneurs engage in crosssector partnerships and bricolage for several other reasons: increased awareness of social problems among the public, coordination of efforts, increased access to financial, human and social capital, enhanced visibility and image, facilitation of social value creation, organisational resilience, and institutionalisation of new approaches and systems (Desa, 2012; Di Domenico, et al., 2010; Korosec & Berman, 2006; Mair & Martí, 2009; Sakarya et al., 2012). Considering the pressures that social enterprises face, collaboration between sectors is essential for achieving scale of impact (Montgomery, Dacin & Dacin, 2012; Sud, VanSandt & Baugous, 2009).

Creating and sustaining social value, impact, or wealth or addressing a social need are usually suggested as the outcomes of social entrepreneurship (Bacq & Jenssen, 2011; Dees, 1998; Mair & Noboa, 2006; Martín & Osberg, 2007). However, social value is a subjective term that varies between geographic and socio-economic contexts (Zahra et al., 2009). What counts as "social" and who needs "help" is a normative judgement. Although most of the research on social entrepreneurship has focused on satisfying basic human needs amongst underprivileged populations, social enterprises can address social problems that affect both disadvantaged population segments and those that might be considered advantaged (Santos, 2012). Thus, the ultimate outcome of socially entrepreneurial efforts is sustainable social transformation, a concept that encompasses narrower terms such as social justice, social wealth or solutions to social issues. It is an outcome that includes different levels of reach and impact based on the fact that social enterprises with different business models in different contexts can transform the cultural, political, and economic lives of millions of people globally or address the very specific needs of a local community (Alvord, et al., 2004).

In summary, social entrepreneurship is the process undertaken by individuals, groups, organisations, and alliances to identify, evaluate, refine, formulate, exploit, and scale opportunities that introduce new goods, services, raw materials, markets, and/or means-ends relations with the primary goal of creating sustainable social transformation locally, regionally, nationally or globally (adapted from Lumpkin et al., 2013; Murphy & Coombes, 2009; Perrini et al., 2010; Shane & Venkataraman, 2000; Zahra et al., 2009). Such a definition recognises social entrepreneurship as a process of creating and organising undertaken by various, and often multiple actors. It positions sustainable social transformation and innovation at the core of the process. It is broad enough to accommodate the numerous manifestations of the process, yet narrow enough to differentiate it from purely social activism, charitable or commercial efforts.

As a special type of entrepreneurship that often results in a hybrid organisation that combines different institutional logics, the process of identifying, evaluating, refining, formulating, exploiting, and scaling social opportunities poses a variety of challenges. Social entrepreneurs design innovative solutions to address complex problems, negotiate, gain and maintain legitimacy, access resources, and create stable and sustainable organisations, alliances and systems in multi-agency environments characterised by uncertainty, ambiguity and complexity. Because of the variety of manifestations of the phenomenon, few guidelines and best practices exist to guide the actions of socially entrepreneurial actors (Lumpkin et al., 2013). As these challenges interact in unpredictable ways, it is not a surprise that nascent and new social entrepreneurship rates are lower than commercial entrepreneurship rates (Lepoutre et al., 2013).

These challenges are most salient in the nascent stages of the process when social opportunities are identified, evaluated, refined, and beginning to be exploited. This is the time when social entrepreneurs design radically new business models to address these individual, organisational and contextual challenges with the aim to create sustainable social transformation. Designing a business model that facilitates creating and sustaining an operational venture and social change in mutually reinforcing ways is a complex endeavour that isn't the result of a single insight, but product of experimentation, adaptation, and refinement (Wilson & Post, 2013; Dees, 1998). Yet very

little research has addressed nascent social entrepreneurship and the processes, mechanisms, behaviours and dynamics involved (Katre & Salipante, 2012; Renko, 2012). A rare contribution in this area is Katre and Salipante's (2012) exploratory investigation of successful and struggling social entrepreneurs' fine-grained behaviours. They suggest that one of the factors that differentiates successful from struggling social entrepreneurs is actively seeking and incorporating feedback to refine and adapt opportunities, products and services.

Although the market itself provides feedback, especially in the form of adversity (Holland & Shepherd, 2013), this is a response to the exploitation of the opportunity. Proactive feedback seeking can minimise risk and reduce uncertainty when applied in the earlier stages of the process to evaluate and refine the opportunity before exploitation. Thus, proactive feedback-seeking behaviour is proposed as a valuable entrepreneurial resource that could help nascent social entrepreneurs address the challenges posed by the process by decreasing uncertainty, minimising risk, setting goals, and increasing innovation, social capital and access to resources.

Proactive feedback seeking and theoretical development

Proactive feedback-seeking behaviour (FSB) is the conscious devotion of effort toward determining the correctness and adequacy of one's behaviour for attaining valued goals (Ashford, 1986). Feedback seeking is grounded in uncertainty reduction theory and allows individuals to gain valuable information about their performance, assess and regulate their performance, gain competence, and correct errors (Ashford & Cummings, 1983).

FSB in the early stages of the entrepreneurial process could reduce uncertainty and minimise risk in the later opportunity exploitation phase by helping entrepreneurs understand the context, the issues, the customers, and the clients and using that information to adapt their ideas. As a valuable uncertainty reduction resource, FSB is often used by individuals in new situations such as organisational entry (Ashford & Black, 1996), career transitions (Callister, Kramer & Turban, 1999), and by young managers (London, Larsen & Thisted, 1999). It is a source of insights and clarity about the environment (Morrison, 1993) and a guide in error correction (Ashford & Cummings, 1983), thus individuals are better able to tailor their behaviour and remain flexible to the unique demands of their individual contexts, which is especially important for nascent social entrepreneurs.

FSB has positive effect on quantity and quality of work through the mediating role of feedback-based goal setting (Renn & Fedor, 2001). Considering that entrepreneurial goals aren't always clear and often generated through improvisation, bricolage or effectuation logics, FSB could be a valuable resource to guide entrepreneurs in setting goals, thus driving the idea in different directions and achieving different ends.

A core characteristic of social entrepreneurship and a risk factor (Renko, 2012) is innovation, which is the implementation of novel and useful solutions to open-ended problems (Amabile, 1997) such as poverty, lack of access to education or healthcare, and social exclusion, among many others. As innovation requires both novelty and usefulness, FSB could be a valuable resource that increases perspective taking and provides new information and ideas. This could lead to the production of novel ideas and evaluation of suitability of ideas for creating stable ventures and sustainable social transformation (Grant & Berry, 2011; Stobbeleir, Ashford & Buyens, 2011).

FSB could enhance the public image and increase social capital of nascent social entrepreneurs, thus affecting legitimacy and access to resources. Since FSB is an inherently social activity, frequency of FSB and diversity of feedback sources could be associated with increased number of relations, as suggested by Apatre and Salipante (2012), and willingness to co-operate. Although there are face costs to FSB, those who seek negative feedback are seen more as attentive and caring (Ashford & Tsui, 1991).

3. Methodology

This research project aims to answer the following research questions:

1. Why, when, and how do nascent social entrepreneurs seek and use feedback?

2. How does FSB relate to opportunity evaluation, refinement, and evolution?

The term opportunity is somewhat misleading here as it implies favourable circumstances and a positive outcome in the form of venture emergence. This research employs a view of opportunities as shaped by actors and expressed as venture ideas that are acted upon without fully knowing whether these ideas will lead to a new venture or whether the new venture will be a direct result of the first business idea (Davidsson, 2003; Dimov, 2011). The entrepreneurship process starts with the conception of a venture idea that includes the the activities and structures that result in value creation. The business idea isn't the product of a single insight, instead it evolves and changes through constant learning, adaptation, and refinement (Wilson & Post, 2013; Dees, 1998), especially in environments of uncertainty (Andries, Debackere & Van Looy, forthcoming).

To adequately answer the two research questions and follow up on the proposed relations between FSB, uncertainty reduction, goal setting, innovation, public image, social capital, and access to resources, a mixed methods approach of in-depth interviews and in situ diary studies is considered. Interviews with nascent, struggling, failed and successful social entrepreneurs recreating "lived worlds" (Spradley, 1979) will be conducted to the point of theoretical saturation (Maxwell, 2005) to gain rich and insightful data on the types of FSB strategies employed and the reasons and circumstances under which nascent social entrepreneurs seek and use feedback. Specific types of feedback-seeking behaviours, data on the patterns of feedback seeking (frequency, sources, source diversity, timing, methods and goals) and idea evolution will be gathered through week-long diary studies every two months for a year, thus limiting retrospective and selection biases (Bolger, Davis & Rafaeli, 2003; Iida et al., 2012) and focusing on the evolution of the venture idea (Davidsson, 2003).

4. Limitations

The proposed research has some conceptual and methodological limitations. Conceptually, the possible negative effects of FSB, such as narrow focus on only one group of stakeholders, conformity, and negative effect on motivation, especially in the case of distractive or low quality feedback, should be considered. The methodological concerns include selection or development of a measure of uncertainty that accurately captures the elements of entrepreneurial uncertainty, adaptation of FSB measures that so far have been used only in cross-sectional studies, controlling for the quality of the feedback, common method bias, access to nascent social entrepreneurs and attrition.

5. Conclusions

The goal of the project is to investigate one of the microfoundations of the social entrepreneurship process and analyse why, when and how nascent social entrepreneurs seek and use feedback and whether and/or how these behaviours are associated with the evaluation and refinement of opportunities and the creation of business ideas that facilitate the creation of stable hybrid ventures and sustainable social transformation. As an inherently social and goal-driven behaviour, FSB enhances both the social and the entrepreneurial in social entrepreneurship. It brings collaboration to the heart of innovation, thus creating value together for all.

Since very little research on nascent social entrepreneurship exists, this project is a step forward for theory and practice, especially with its focus on idea evolution and design of hybrid models from the beginning of the entrepreneurial process. The research directly contributes to the literatures on nascent social entrepreneurship and entrepreneurial opportunities, while generating knowledge useful for and directly applicable to practitioners, including educators and staff members in incubators, accelerators and development agencies. The project will extend the concept of FSB to a new domain. In addition to being an organisational and individual resource, FSB is also proposed as an entrepreneurial resource.

References

- Alvord, S.H., Brown, L.D., & Letts., C.W. (2004). Social entrepreneurship and social transformation: An exploratory study. *The Journal of Applied Behavioral Science 40*, 3, p. 260–82.
- Amabile, T.M. (1997). Entrepreneurial creativity through motivational synergy. Journal of Creative Behavior, 31, 1, p. 18-26.
- Andries, P., Debackere, K., & Van Looy, B. (forthcoming). Simultaneous experimentation as a learning strategy: Business model development under uncertainty. *Strategic Entrepreneurship Journal.*
- Ashford, S.J. (1986). The role of feedback seeking in individual adaptation: A resource perspective. *Academy of Management Journal, 29*, p. 465-487.
- Ashford, S.J., & Black, J.S. (1996). Proactivity during organizational entry: The role of desire for control. *Journal of Applied Psychology*, *81*, p. 199-214.
- Ashford, S.J., & Cummings, L.L. (1983). Feedback as an individual resource: Personal strategies for creating information. Organizational Behavior and Human Performance, 32, p. 370-398.
- Ashford, S.J., & Tsui, A.S. (1991). Self-regulation for managerial effectiveness: The role of active feedback seeking. *Academy of Management Journal*, 34, p. 251-280.
- Austin, J., Stevenson, H., & Wei-Skillern, J. (2006). Social and commercial entrepreneurship: Same, different, or both? *Entrepreneurship Theory and Practice 31*, 1, p. 1–22.

- Bacq, S., Janssen, F. (2011). The multiple faces of social entrepreneurship: A review of definitional issues based on geographical and thematic criteria. *Entrepreneurship & Regional Development*, 23, 5/6, p. 73-403.
- Battilana, J. & S. Dorado. (2010). Organization building amid multiple institutional logics: The case of commercial microfinance organizations. Academy of Management Journal, 53, 6, p. 1419-1440.
- Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology, 54*, p. 579-616.
- Bornstein, D. (1998). Changing the world on a shoestring. *The Atlantic Monthly 281*, 1, p. 34–39.
- Callister, R.R., Kramer, M.W., & Turban, D.B. (1999). Feedback seeking following career transitions. *Academy of Management Journal*, 42, 2, p. 429-438.
- Davidsson, P. (2003). The domain of entrepreneurship research: Some suggestions. In J.A. Katz & D.A. Shepherd (Eds.), *Advances in entrepreneurship, firm emergence and growth* (Vol. 6, p.315–372). Oxford, UK: Elsevier/JAI Press.
- Dees, G. (1998). The meaning of social entrepreneurship. Durham, NC: Center for the Advancement of Social Entrepreneurship, Duke Faqua School of Business. Retrieved on April 18, 2013 from http://caseatduke.org/documents/dees_ sedef.pdf.
- Desa, G. (2012). Resource mobilization in international social entrepreneurship: Bricolage as a mechanism of institutional transformation. *Entrepreneurship Theory and Practice*, 36, 4, p. 727-751.
- Di Domenico, H., Haugh, H., & Tracey, P. (2010) Social bricolage: theorizing social value creation in social enterprise. *Entrepreneurship Theory and Practice*, 34, 4, p. 681-703.
- Dimov, D. 2011. Grappling with the unbearable elusiveness of entrepreneurial opportunities. *Entrepreneurship Theory and Practice*, 35, 1, p. 57-81.
- Galaskiewicz, J., & Barringer, S.N. (2012). Social enterprises and social categories. In B. Gidron & Y. Hasenfield (Eds.), *Social Enterprises: An Organizational Perspective* (p. 47-90). Houndmills: Palgrave Macmillan.
- Grant, A.M. & Berry, J.W. (2011). The necessity of others is the mother of invention: intrinsic and prosocial motivations, perspective taking, and creativity. *Academy of Management Review, 54,* 1, p. 73-96.
- Holland, D.V., & Shepherd, D.A. (2013). Deciding to persist: Adversity, values, and entrepreneurs' decision policies. *Entrepreneurship Theory and Practice*, *37*, 2, p. 331-358.
- Iida, M., Shrout, P.E., Laurenceau, J.-P., & Bolger, N. (2012). Using diary methods in psychological research. In H. Cooper (Ed.), APA handbook of research methods in psychology, Vol. 1: Foundations, planning, measures, and psychometrics (p. 277-305). Washington, DC: American Psychological Association.
- Katre, A., & Sapilante, P. (2012). Start-up social ventures: Blending fine-grained behaviors from two institutions for entrepreneurial success. *Entrepreneurship Theory and Practice, 36, 5, p. 967-994.*
- Koellinger, P. (2008). Why are some entrepreneurs more innovative than others? *Small Business Economics*, *31*, 1, p. 21–37.

- Korosec, R.L., & Berman, E.M. (2006). Municipal social support for social entrepreneurship. *Public Administration Review*, 66, 3, p. 448-462.
- Lepoutre, J., Justo, R., Terjesen, S., & Bosma, N. (2013). Designing a global standardized methodology for measuring social entrepreneurship activity: The Global Entrepreneurship Monitor social entrepreneurship study. *Small Business Economics*, 40, 3, p. 693-714.
- London, M., Larsen, H.H., & Thisted, L.N. (1999). Relationships between feedback and self-development. *Group & Organization Management*, 24, 5, p. 5-26.
- Lumpkin, G.T., Moss, T.W., Gras, D.M., Kato, S., & Amezcua, A.S. (2013). Entrepreneurial processes in social contexts: How are they different, if at all? *Small Business Economics*, 40, 3, p. 761-783.
- Mair, J., Battilana, J., & Cardenas, J. (2012). Organizing for society: A typology of social entrepreneuring models. *Journal of Business Ethics*, 111, 3, p. 353-373.
- Mair, J., & Martí, I. (2006). Social entrepreneurship research: A source of explanation, prediction and delight. *Journal of World Business* 41, 1, p. 36– 44.
- Mair, J., & Martí., I. (2009). Entrepreneurship in and around institutional voids: A case study from Bangladesh. *Journal of Business Venturing 24*, 5, p. 419– 435.
- Mair, J., Martí, I., &Ventresca, M. (2012). Building inclusive markets in rural Bangladesh: How intermediaries work institutional voids. Academy of Management Journal, 55, 4, p. 819-850.
- Mair, J., & Noboa, E. (2006). Social entrepreneurship: How intentions to create a social venture are formed. In J. Mair, J.A. Robinson, & K. Hockerts (Eds.), *Social entrepreneurship* (p. 121–35). Basingstoke and New York: Palgrave Macmillan.
- Mair, J., & Schoen, O. 2007. Social entrepreneurial business models in the context of developing economies: An exploratory study. *International Journal of Emerging Markets*, 2, 1, p. 54-68.
- Martin, R.L., & Osberg, S. (2007). Social entrepreneurship: The case for definition. *Stanford Social Innovation Review*, Spring, p. 29–39.
- Maxwell, J.A. (2005). *Qualitative research design: An interactive approach* (2nd edition). Thousand Oaks, CA: Sage Publications.
- Moizer, J., & Tracey, P. (2010). Strategy making in social enterprises: The role of resource allocation and its effects on organizational sustainability. *Systems Research and Behavioral Science*, 27, 3, p. 252-266.
- Montgomery, W., Dacin, P.A., & Dacin, M.T. (2012). Collective social entrepreneurship: Collaboratively shaping social good. *Journal of Business Ethics*, 111, 3, p. 375-388.
- Morrison, E. (1993). Longitudinal study of the effects of information seeking on new comer socialization. *Journal of Applied Psychology*, 78, p.173-183.
- Murphy, P.J., & Coombes, S.M. (2009). A model of social entrepreneurial discovery. *Journal of Business Ethics*, 87, 3, p. 325-336.
- PBS. (2005). The new heroes. Who were the first social entrepreneurs? Retrieved on 13.02.2013 from http://pbs.org/opb/thenewheroes/whatis/whatis_ss_6.html.

- Peredo, A.M., & McLean, M. (2006). Social entrepreneurship: A critical review of the concept. *Journal of World Business* 41,1, p. 56–65.
- Perrini, F., Vurro, C., & Costanzo, L.A. (2010). A process-based view of social entrepreneurship: From opportunity identification to scaling-up social change in the case of San Patrignano. *Entrepreneurship & Regional Development, 22,* 6, p. 515-534.
- Renko, M. (2012), Early Challenges of Nascent Social Entrepreneurs. Entrepreneurship Theory and Practice. doi: 10.1111/j.1540-6520.2012.00522.x
- Renn, R.W., & Fedor, D.B. (2001). Development and field test of a feedback seeking, self-efficacy, and goal setting model of work performance. *Journal of Management, 27*, p. 563-584.
- Sakarya, S., Bodur, M., Yildirim-Öktem, O., & Selekler-<u>Göksen</u>, N. (2012). Social alliances: Business and social enterprise collaboration for social transformation. *Journal of Business Research*, 65, 12, p. 1710-1720.
- Samuelsson, M. & Davidsson, P. (2009). Does venture opportunity variation matter? Investigating systematic process differences between innovative and imitative new ventures. *Small Business Economics*, 33, 2, p. 229–255.
- Santos, F. (2012). A positive theory of social entrepreneurship. *Journal of Business Ethics*, 111, 3, p. 335-351.
- Schumpeter, J.A. (1961). *The theory of economic development: An inquiry into profits, capital, credit, interest and the business cycle.* Cambridge, MA: Harvard University Press.
- Seelos, C., & Mair, J. (2005). Social entrepreneurship: Creating new business models to serve the poor. *Business Horizons 48*, 3, p. 241–246.
- Seelos, C., & Mair, J. (2007). Profitable business models and market creation in the context of deep poverty: A strategic view. Academy of Management Perspectives 21, 4, p. 49–63.
- Seelos, C., Mair, J., Battilana, J., & Dacin, M.T. (2011). The embeddedness of social entrepreneurship: Understanding variation across geographic communities. In C. Marquis, M. Lounsbury, & R. Greenwood (Eds.), *Research in the Sociology of Organizations. Communities and Organizations* (Vol. 33, p. 333-363). Emerald Group Publishing Limited.
- Shane, S., & Venkataraman, S. (2000). The promise of entrepreneurship as a field of research. *Academy of Management Review 25*, 1, p. 217–26.
- Shaw, E. & Carter, S. (2007). Social entrepreneurship: Theoretical antecedents and empirical analysis of entrepreneurial processes and outcomes. *Journal of Small Business and Enterprise Development*, 14, 3, p. 418-434.
- Shepherd, D.A., Douglas, E.A., & Shanley, M. (2000). New venture survival: Ignorance, external shocks, and risk reduction strategies. *Journal of Business Venturing*, 15, p. 393–410.
- Smallbone, D., Evans, M., Ekanem, I., & Butters, S. (2001). Researching social enterprise: Final report to the small business service. London: Centre for Enterprise and Economic Development Research, Middlesex University.
- Spradley, J. (1979). *The ethnographic interview*. New York, NY: Holt Rinehart and Winton.

- Stobbeleir, K.E.M., Ashford, S.J., & Buyens, D. (2011). Self-regulation of creativity at work: The role of feedback-seeking behavior in creative performance. *Academy of Management Journal*, 54, 4, p. 811-831.
- Stryjan, Y. (2006). The practice of social entrepreneurship: Notes toward a resource perspective. In C. Steyart & Hjorth, D. (Eds.), *Entrepreneurship as social change: A third movements in entrepreneurship book* (p. 35–55). Cheltenham and Northampton: Edward Elgar
- Suchman, M.C. (1995). Managing legitimacy: Strategic and institutional approaches. Academy of Management Review, 20, 3, p. 571-610.
- Sullivan Mort, G., Weerawardena, J., & Carnegie, K. (2003). Social entrepreneurship: Towards conceptualization. *International Journal of Non-profit and Voluntary Sector Marketing*, 8, 1, p. 76–88.
- Sud, M., VanSandt, C.V., & Baugous, A.M. (2009). Social entrepreneurship: The role of institutions. *Journal of Business Ethics*, 85, 1, p. 201-216.
- Thompson, J.L., Alvy, G., & Lees, A. (2000). Social entrepreneurship A new look at the people and the potential. *Management Decision 38*, 5, p. 328–338.
- Tracey, P., & Phillips, N. (2007). The distinctive challenge of educating social entrepreneurs: A postscript and rejoinder to the special issue on entrepreneurship education. *Academy of Management Learning and Education 6, 2, p. 264–271.*
- VanSandt, C.V., Sud, M., & Marme, C. (2009). Enabling the original intent: Catalyst for social entrepreneurship. *Journal of Business Ethics*, *90*, 3, p. 419-428.
- Weerawardena, J., & Sullivan Mort, G. (2006). Investigating social entrepreneurship: A multidimensional model. *Journal of World Business 41*, 1, p. 21–35.
- Weerawardena, J., & Sullivan Mort, G. (2012). Competitive strategy in socially entrepreneurial nonprofit organizations: Innovation and differentiation. *Journal of Public Policy & Marketing*, 31, 1, p. 91-101.
- Wilson, F., & Post, J.E. (2013). Business models for people, planet (& profits): Exploring the phenomena of social business, a market-based approach to social value creation. *Small Business Economics*, 40, p. 715-737.
- Yunus, M. (2007). Creating a world without poverty: Social business and the future of capitalism. New York, NY: PublicAffairs.
- Zahra, S.A., Gedajlovic, E., Neubaum, D.O., & Shulman, J.M. (2009). A typology of social entrepreneurs: Motives, search processes and ethical challenges. *Journal of Business Venturing*, 24, 5, p. 519-532.

Investigating the causality between Financial Development and Economic Growth in the developing countries of Europe: Evidence from Albania and Turkey

Esida Gila Gourgoura¹, Eftychia Nikolaidou²

¹ Business Administration and Economics Department, CITY College - International Faculty of the University of Sheffield, egila@city.academic.gr

² Business Administration and Economics Department, CITY College - International Faculty of the University of Sheffield, enikolaidou@city.academic.gr

Abstract. This paper seeks to investigate the causal relationship between financial development and economic growth for Albania and Turkey, over the 1998-2012 period. The international academic literature has not reached a consensus regarding the finance-growth issue, whereas particularly in the two countries, relevant empirical work is limited. For the purpose of investigation, five proxies of financial development are introduced in the case of Albania whereas considering its developed financial market two more indicators are added in the case of Turkey. The real GDP is used as proxy for economic growth. As soon as cointegration is established between each indicator of finance and the real GDP, it is tested for causality through the Vector Error Correction Model. Findings suggest that bi-directional causality exists between financial development and economic growth in Albania, both in the short and in the long-run. The same feedback relationship is discovered for Turkey but only in the long-run. Apart from the causality findings, it is noted that two indicators of the banking sector development have a significant negative impact on growth, more precisely, loans to deposits ratio and bank capital to assets ratio. To this extent, our findings have certain implications for the regulation of the banking sector in the countries under investigation and those similar to them.

Keywords: Financial development, economic growth, causality, developing countries

1. Introduction

Economic growth and other issues related to it have drawn scholars' attention since the early 1900s. Obviously, the main concern has been establishing which are the potential factors that affect growth. Generally speaking, the increase of a country's productive capacity over time, in other words growth, can be achieved due to an efficient capital accumulation and its allocation to the most productive sectors. Considering that this comprises one of the main functions of the financial intermediaries it may be concluded that a positive relationship exists between finance and growth.

Indeed, Levine (2004) argues that by performing efficiently their functions¹, financial intermediaries can essentially enhance the overall capacity of the economy. One should note though, that the soundness of the financial system is crucial to its role as growth accelerator. The opposite may occur if as finance develops, certain shortcomings are identified in its activity that may cause financial crises and hit the economy (Khodzhibaeva and Nikolaidou, 2012). Therefore, the finance-growth nexus has been the focus of a vast amount of academic studies. Still, the evidence is far from conclusive.

This paper empirically investigates the finance-growth causality in two developing countries of Europe, namely, Albania and Turkey. The rationale of our choice is related to the insufficient existing empirical work, especially in Albania. Moreover, Albania and Turkey represent two countries, equally classified as developing economies (IMF, 2013) and yet far different in terms of their financial system's composition and economic development stage. To this extent, comparing the results between the two may provide a more complete view on the finance-growth link among various stages of development of the emerging economies.

The paper is organized as follows: Section 2 discusses the academic literature on the finance-growth nexus whereas Section 3 provides a brief overview of the Albanian and Turkish financial and economic development. Section 4 describes the data used and the methodology followed for empirical investigation whereas Section 5 examines the empirical findings. The paper concludes with Section 6.

2. Literature Review

Levine (1997) highlights the two main theoretical views on the relationship between financial development and economic growth. The first one which is known as the neoclassical growth theory is focused on the capital allocation function of the financial intermediaries and explains how it affects growth (i.e. by transferring funds among different industries or by changing savings rate). The second theory deal with productivity thus is focused on other functions of financial sector that improve productivity or introduce in the market new innovative products. Demetriades and Andrianova (2004) though, criticized the neoclassical theory assumptions of the "neu-

¹ i) asset transformation function (including: maturity transformation, diversification and management of risk, reduction of transaction costs), ii) facilitating the payments mechanism in the exchange of goods and services and iii) risk management activities

tral" role of finance in economic growth and the existence of "perfect capital markets". The authors argued that the reality is far from perfection, since asymmetric information and transaction costs exist. Therefore, by reducing these two threats, financial intermediaries play an important role in improving the quality of investments, thus helping economic growth.

Patrick (1966) introduced the "supply-leading" and "demand-following" views on the finance-growth link. According to the first one, in the very early stages of development, causality often runs from economic to financial development. In other words, financial development follows economic growth as a result of an increased demand for financial services. As economic growth occurs though, the direction of causality may reverse and a "supply-leading" relationship develops. Here, the efficiency gains associated with the intermediation process help generate continued economic growth. To this extent, bi-directional causality may occur between financial development and economic growth.

Empirical models that examine the relationship between financial development and economic growth evolve from earlier formal models (i.e. neoclassical growth models) that simply test the finance – growth nexus, to recent ones that aim to examine the direction of causality between them. Moreover, as finance develops and the complexity of its products and services increases, the variables that proxy it in the studies change over time.

Demetriades and Law (2006) employed a formal neoclassical model (Cobb-Douglas production function) to examine the importance of the financial sector's quality in the economic growth of 72 countries over the period 1978–2000. The authors concluded that the latter is affected by both the quantity and the quality of the financial sector's services. King and Levine (1993) performed a cross-section regression in 80 countries over the 1960-1989 period in order to define the finance-growth relationship. Findings suggested that financial development has a positive effect on economic growth. In contrast, Rioja and Valev (2004) failed to find such relationship when various low- and middle-income countries were analysed for the period 1961-1995. Their findings suggest that only in high-income countries financial development significantly affects economic growth. Atakan et al (2010) investigated the impact of the financial sector development for the E-7 countries over the period 2001-2007. They utilized eight indicators in their panel regression that resemble growth and concluded that five of them are highly significant for growth.

Levine and Zervos (1998) studied the effects of both banks and financial markets in economic growth, thus, introduced additional indicators of financial development. After an assessment of 47 countries for the period 1976-1993, their findings suggested that banks and markets are complements in affecting positively economic growth. Beck and Levine (2001) combined both cross-border regression and time series analysis in a panel of 40 countries and 146 observations. They concluded that both stock markets and banks have a positive impact on growth. Liang et al (2007) investigated a large group of emerging markets/developing countries as well as advanced economies over the 1998-2003 period. In contrast to Levine and Zervos they found an important shift from a heavy reliance on basic banking services among both developing and advanced countries towards an expanded role for the capital markets. This shift was more noticeable among the advanced countries.

Rosseau and Wachtel (1998) concluded in a long- run relationship between finance and growth when the link was analyzed for 5 industrialized countries (United States, the United Kingdom, Canada, Norway, and Sweden) over the 1870-1929 period. The Granger causality test's results suggested that are the intermediaries that lead to economic growth but failed to find feedback relationship between them. Calderon and Liu (2003) examined the direction of causality between financial development and economic growth of 109 developing and industrial countries from 1960 to 1994. The results suggested bi-directional relationship between the two variables for all the countries. Hondroyiannis et al (2005) investigated the relationship between the development of the banking system and the stock market and economic performance in Greece over the period 1986-1999 (monthly data). Findings suggested that both the indicators of finance can promote growth in the long- run but the impact of the stock market is rather small.

Recent academic research attempts to investigate developing countries as well. Results from the empirical work of Saibu et al (2011) suggest that there is unidirectional causality from financial development to economic growth in Nigeria. The authors included also foreign direct investment in their study as a control variable and found that also the latter causes economic growth. Hosny (2012) conducted the same study for Egypt for the period 1961-2009. He concluded that finance Granger causes economic growth in both the short and the long-run, thus providing support for the supply leading hypothesis. Dushku (2010) investigated the finance-growth link for Albania over the period 1996-2007. Findings suggested that there is a bi-directional relationship between financial development and economic growth, in the long run. In the short term though, the relationship could not be clearly established. Consistently, bi-directional causality (but both in the short and in the long term) was achieved from Demirhan et al (2011) when the same investigation was performed for Turkey over the period 1987:1-2006:04. In contrast, Soytas and Kucukkaya (2011) failed to find causality relationship between finance and growth in Turkey for the period 1991:3-2005:4. Their findings suggested that financial development and economic growth cannot improve each others' forecasts.

3. Albania and Turkey: Financial Development and Economic Growth

3.1. Albania

According to the Bank of Albania (BoA) stability report of the first half of year 2012, banking sector has always been the dominant segment of financial intermediation in Albania. Prior to the collapse of the communist regime in the early 1990s, the Albanian banking system was entirely state-owned (Dushku, 2010). The existence of a two-tier banking system was approved by the new Banking Law of the year 1992. This year encountered also structural economic reforms in the country such as the privatization of the state-owned premises and the establishment of a free market economy.

To this extent, the economy entered the growth path, more precisely, the growth rate increased to 8% and 9% in the years 1995 and 1996 (Jarvis, 2000).

Meanwhile, the infusion of foreign financial institutions in the domestic banking sector, mainly Greek and Italian ones, expanded the variety of banks operating in Albania to 8. The collapse of the Pyramid schemes though in the year 1997, associated with a massive rebellion of the Albanian citizens caused huge damages to the social stability of Albania (Jarvis, 2000). Currently, there are 16 banks operating in Albania, whose assets accounted for 82% of the GDP in the first half of the year 2012 (BoA, 2012). Consumer lending is the main focus of their activity in the last decade whereas there is lack of other basic banking activities such as commercial and investment banking. The facts that no stock exchange exists in Albania and its banking system has a limited activity indicate that the financial development in Albania is still in its earliest stage (Dushku, 2010). Based on the BoA stability report (2011 H2), the annual real GDP growth rate in Albania started its ascending trend after the year 2004, reaching to 5%, 6% and 7.7% in the years 2006, 2007 and 2008. Moreover, statistics suggest that Albania featured the higher growth rate in the South-East Europe (Dushku, 2010). However, according to the recent stability report of BoA (2012), the real growth rate in Albania deteriorated to 3.5% in the year 2010 and 3% in the year 2011 and it was associated with higher unemployment and inflation rates.

3.2. Turkey

Until the early1980s, Turkey's financial sector was highly restricted (Kar et al, 2008). Its liberalization initiated when certain financial organizations were established. The Capital Markets Board of Turkey (CMB) founded in the year 1982 encouraged open market operations whereas the Banking Law of 1985 enacted interbank activities (Demirhan et al, 2011). Furthermore, the Istanbul Stock Exchange was established in 1986 followed by the establishment of the Gold Stock Exchange in 1989. In parallel with the financial system's development, Turkey's economy entered its liberalization path. Significant structural changes occurred, however, due to the weaknesses of the undertaken reforms, instability still characterized the economy (Disbudak, 2010).

Extremely high inflation rates (40%) were registered during the 1980s and the 1990s due to continuous money printing from the Central Bank. Furthermore, the economy of Turkey suffered from a number of periodic crises such as the ones of the years 1994 and 2001. Inevitably, the banking sector encountered problems as well; 5 banks were closed during the 1990s whereas the partnership rights of 20 banks were transferred to The Savings Deposit Insurance Fund (Demirhan et al, 2011). The tough financial condition urged the need for the Supervisory and Regulatory Board of Banking, which was established in the year 2000 and played a significant role in the country's financial upturn. Furthermore, new policies were developed which increased the efficiency of financial intermediation during the last decade (Kar et al, 2008).

Currently, according to the National Bank of the Republic of Turkey (NBRT) stability report of the year 2012, there are 48 banks operating in Turkey, whose assets accounted for 93% of the GDP in the end of June 2012. In contrast to the Albanian banking system, the Turkish one consists of deposit banks, investment and participation banks. To this extent, it is considered as one of the most developed banking systems in East Europe and Central Asia. In parallel with a sound banking system, a rapid economic growth is witnessed in Turkey. More precisely, the economy achieved high levels of growth in the last years (8.5% in the second half of year 2011), ranking as the 15th in the world with the largest GDP (NBRT, 2012). In the first half of the year 2012, GDP rose 4.9% in annual terms.

4. Data and Methodology

In order to perform the investigation, the following indicators of financial development are used for both countries: bank credit to private sector as a share of GDP, bank deposits as a share of GDP, bank capital to assets ratio, loans to deposits ratio and broad money (M2) as a share of GDP. In the case of Turkey, two additional indicators are used, which proxy stock market development such as market capitalization as a share of GDP and the value of stocks traded in the Istanbul Stock Exchange as a share of GDP. Economic growth is indicated by the real GDP.

Considering that the banking system represents the main body of a country's financial system, it becomes obvious that the efficiency and the quality of its activity reflect the development of the financial system as a whole (King and Levine, 1993). Bank credit to private sector as a share of GDP is commonly used in the relevant literature as one of the main indicators of financial development. Indeed, continuous investments in the private sector may boost the economy. Moreover, banks are quite strict on evaluating the quality of the projects that will be financed by them, thus, through credit to private sector the possibility for sound investments is higher (Levine and Zervos, 1998).

According to Levine (2004), by pooling the savings of individuals, banks engage in enormous financial arrangements. For this to occur though, they have to convince savers of the soundness of the investments. In other words, the bigger the bank deposits, the higher the credibility of a bank, the better the quality of its investments. Therefore, bank deposits are considered as another indicator of financial development and are supposed to have a positive effect on growth.

Liquidity risk is one of the most discussed issues related to the recent "financial deepening" of the economic systems. More precisely, in the last decade, banks have increased the volume of loans not only at a faster rate than economic activity but also more rapidly than the liquidity that they create in the form of deposits (Gualandri et al, 2009). Consequently, liquidity risk has emerged simultaneously with the development of the financial systems. For this purpose, loans to deposits ratio is used in this study as to indicate an important development of the financial systems. Its effect on economic growth though, may not necessarily be positive, considering the liquidity problems related to it.

Banks operate with high financial leverage and their financial risk has been fostered especially in emerging countries during crises periods. This has caused many banks go in bankruptcy which renders banks' capital a vital issue for healthiness of the banking sector. Thus, bank capital to asset ratio (CAR hereafter) has been used as a proxy for measuring weaknesses against high financial risk (Atakan et al, 2010). One should note though, that a low level of leverage means not enough sound investments undertaken from the banks, thus, its effect on growth may also be negative.

Broad money (M2) as a share of GDP is considered the most appropriate monetary aggregate to proxy financial development. A high level of monetization (M1/GDP) is most likely the result of financial underdevelopment whereas M3/GDP depends rather on other indicators than on financial development (Dushku, 2010).

The development of financial markets drives financial innovation, efficiency, and technological advancement (Zhou, 2011). Therefore, two indicators of the financial markets such as market capitalization and the value of the stocks traded in the ISE as a share of GDP are included as measures of financial development in our study. The market capitalization gives the total value of all listed companies in the stock market and may be considered as a good indicator for the stock market size and ability of companies to raise equity funds in economies (Atakan et al, 2010). On the other hand, the value of the stocks traded in the ISE gives a robust activity of the stock markets. It is worth noting that these indicators are only included in the case of Turkey since no stock exchange exists in Albania.

With reference to the main body of literature, the focus of our analysis is constructing the underlying VAR model that consists of a set of endogenous variables. VAR models are mostly used to test Granger causality and to determine whether lags of one variable help explain the current value of some other variables (Granger, 1969).

$$\Delta EG = \alpha_0 + \alpha_1 \Delta FD_{t-1} + \alpha_2 \Delta FD_{t-2} + \alpha_3 \Delta EG_{t-1} + \alpha_4 \Delta EG_{t-2} + \alpha_5 ECT_t$$

 $\Delta FD = \beta_0 + \beta_1 \Delta EG_{t-1} + \beta_2 \Delta EG_{t-2} + \beta_3 \Delta FD_{t-1} + \beta_4 \Delta FD_{t-2} + \beta_5 ECT_t$

where,

 $\triangle EG$ and $\triangle FD$ are endogenous variables,

 α_0 and β_0 are the constants,

 ΔEG_{t-1} , ΔEG_{t-2} and ΔFD_{t-1} , ΔFD_{t-2} are the lagged value of the endogenous variables for one and two periods respectively,

ECT is the error correction term (the lagged residuals)

To obtain stable results, the series tested for Granger causality should be stationary. Since most time series variables are non-stationary, unit root tests are useful to determine the order of integration of the variables. For this purpose, Augmented Dickey-Fuller (ADF) test for the existence of any unit root will be applied.

As soon as all the variables are transformed into stationary, it will be tested whether cointegration (long-run relationship) exists between them, through the Engle and Granger two step approach. Many studies base their cointegrating tests on the Johansen's maximum likelihood approach. However, the latter is more appropriate in a more than 2 variable model since it can estimate more than one cointegration relationship (Gujarati, 2009). Engle and Granger (1987) pointed out that a linear combination of two or more non-stationary series may be stationary. If such a stationary linear combination exists, the non-stationary variables are said to be cointegrated, that is, there is a long-run relationship between them. The Engle and Granger two step approach for cointegration is as follows: Firstly, we run the OLS regression on what is called the cointegrating equation: $EG_t = \alpha_0 + \alpha_1 FD_t + e_t$. Secondly, we test whether the residuals e_t from the regression are stationary. If the residuals are found stationary, that is I(0), then a long-run relationship exists between the variables.

In our case, the existence of cointegration among each indicator of financial development and the growth's indicator implies that at least one causes the other (Engle and Granger, 1987). Therefore, there are two possible sources of causality: error correction term, which shows long-run causality, and lagged explanatory variables, revealing short-term causality (Demirhan et al, 2011). Error correction term indicates the short run deviations from the long run equilibrium and its size shows the speed of adjustment of any disequilibrium towards long run equilibrium (Granger, 1969). In addition, Wald statistics are useful in determining the short-term causality. To determine Granger causality from financial development to economic growth, the Wald test statistics are calculated under the null hypothesis that all coefficients of $\alpha_1... \alpha_4=0$ as a group. Similarly, applying Wald test for causality from growth to finance requires testing the null hypothesis that all coefficients of $\beta_1... \beta_4 = 0$ as a group. Rejecting the null hypothesis that the explanatory variables as a group do not affect the particular dependent variable indicates short-term Granger causality from explanatory variables to the latter (Humpage and Pelz, 2002).

The variables used in this study are as following: *LGDPAL* and *LGDPTR* are the real GDP for Albania and Turkey; *DEPAL* and *DEPOSITTR* are the bank deposits as a ratio of GDP for Albania and Turkey; *PRCRAL* and *PRIV_CREDITTR* are the bank credit as a ratio of GDP for Albania and Turkey; *LOANS_DEPAL* and *LOANS_DEPTR* are the loans to deposits ratios for Albania and Turkey; *M2AL* and *M2TR* are the M2 to GDP ratios for Albania and Turkey; *CARAL* and *CARTR* are the bank capital to assets ratios for Albania and Turkey; *MARK_CAP* is the market capitalization as a ratio to GDP for Turkey; *STOCKS_TRD* is the total value of the stocks traded in the Istanbul Stock exchange as a ratio of GDP.

All the variables are transformed into their natural logarithms to avoid the problem of heteroskedasticity. Data for all indicators are sourced from the official sites of their respective Central Banks. Moreover, INSTAT (Statistics Institute) is another source of data for Albania. Quarterly observations are used in our investigation that span from the fourth quarter of the year 1998 to the fourth quarter of the year 2012.

5. Empirical Findings

5.1. Unit Root Test

Being stationary is conditional for performing the Granger causality tests, thus, the Augmented Dickey-Fuller (ADF hereafter) test will be applied for each variable in two cases: with the intercept and both with the intercept and the trend. Table 1 repre-

sents the unit root test results for Albania whereas the results for Turkey are presented in Table 2.

	Le	evel	First Di	ifference	Second Difference		
Indicators	Intercept	Intercept & trend	Intercept	Intercept & trend	Intercept	Intercept & trend	
LGDPAL	-2.3(3)	-1.73(4)	-18.6(2)***	-19.4(2)***			
DEPAL	0.48(4)	-2.41(4)	-3.43(3)**	-3.6(3)**			
PRCRAL	-0.63(4)	-2.07(4)	-1.75(3)	-2.75(3)	-27(2)***	-27(2)***	
LOANS_DEPAL	-0.56(4)	-2.59(4)	-1.72(3)	-1.55(3)	-23(2)***	-23(2)***	
M2AL	-2.16(3)	-2.19(3)	-21.6(2)***	-21.5(2)***			
CARAL	-1.63(3)	-3.6(3)**					

Table 1. ADF Unit Root Test results for Albania

** and *** denote stationary at 5% and 1% level of significance.

The 10%, 5% and 1% critical values are -2.59, -2.92 and -3.56 for the ADF test with the Intercept and -3.18, - 3.50 and -4.15 for the one with the Intercept and the Trend. The proper lag order for ADF test is chosen by considering (AIC) and white noise of residuals, representing in parenthesis. As it is noted in the above table, CARAL is an I(0) variable since it is stationary in its level, LGDPAL, DEPAL and M2AL become stationary in their first differences, whereas PRCRAL and LOANS_DEPAL are I(2), thus, integrated of order two series.

	L	evel	First I	Difference
Indicators	Intercept	Intercept & trend	Intercept	Intercept & trend
LGDPTR	-0.79(10)	-1.67(10)	-2.7(2)*	-3.2(3)*
DEPOSITTR	1.09(5)	-1.18(5)	-3.27(4)**	-3.57(4)**
PRIV_CREDITTR	3.64(6)	0.69(9)	-2.61(4)*	-4.37(5)***
LOANS_DEPTR	3.71(6)	1.03(6)	-2.72(7)*	-4.51(5)***
M2TR	069(5)	-2.1(5)	-3.3(4)**	-3.52(4)**
CARTR	-2.25(5)**	-1.22(4)	-3.24(4)**	-3.52(4)**
MARK_CAP	-0.89(0)	-3.41(1)*		
STOCKS_TRD	-1.52(1)	-3.77(5)**		

 Table 2. ADF Unit Root Test results for Turkey

*, ** and *** denote stationary at 10%, 5% and 1% level of significance.

The 10%, 5% and 1% critical values are -2.59, -2.92 and -3.56 for the ADF test with the Intercept and -3.18, - 3.50 and -4.15 for the one with the Intercept and the Trend. The proper lag order for ADF test is chosen by considering (AIC) and white noise of residuals, representing in parenthesis. As it is noted in Table 2, most of the

variables become stationary at their first differences, despite MARK_CAP and STOCKS_TRD, which are I(0) series, thus, stationary in their level.

5.2. Cointegration Test

When two series are assumed to be non-stationary, a cointegration test can be applied to check for a long-run relationship between the variables. For this purpose, the Engle and Granger two-step approach is applied in our study. The results for Albania are shown in Table 3.

Cointegration regression	R-squared	ADF test statistic			
		Intercept	Intercept & trend		
LGDPAL=10.75+0.64*DEPAL	0.62	-6.3***	-7.54***		
LGDPAL=11.67-0.44*LOANS_DEPAL	0.79	-3.49**	-3.19**		
LGDPAL=11.68+0.44*PRCRAL	0.83	-5.01***	-5.09***		
LGDPAL=11.63+0.18*M2AL	0.78	-4.03***	-9.43***		
LGDPAL=12.24-15.4*CARAL	0.16	-3.05**	-5.32***		

Table 3. Engle-Granger approach results for Albania

** and *** denote stationary of the residuals at 5% and 1% level of significance.

The 10%, 5% and 1% critical values are -2.59, -2.92 and -3.56 for the ADF test with the Intercept and -3.18, - 3.50 and -4.15 for the one with the Intercept and the Trend. As observed in the above table, the residuals from all the regressions are stationary, thus, it is concluded that cointegration exists between each indicator of financial development and economic growth in Albania. Results for Turkey are outlined in Table 4.

Cointegration regression	R-squared	ADF	test statistic
		Intercept	Intercept &trend
LGDPTR=16.63+0.02*DEPTR	0.63	-7.16***	-7.38***
LGDPTR=16.71-0.02*LOANS_DEPTR	0.63	-6.28***	-6.45***
LGDPTR=16.71+0.03*PRIV_CREDITTR	0.63	-6.3***	-6.48***
LGDPTR=16.69+0.02*M2TR	0.66	-7.74***	-7.87***
LGDPTR=16.4+4.09*CARTR	0.34	-3.03**	-4.42***
LGDPTR=16.64+32.38*MARK_CAP	0.68	-8.5***	-9.12***
LGDPTR=16.72+3.67*STOCKS_TRD	0.38	-4.61***	-6.05***

Table 4. Engle-Granger approach results for Turkey

** and *** denote stationary of the residuals at 5% and 1% level of significance.

The 10%, 5% and 1% critical values are -2.59, -2.92 and -3.56 for the ADF test with the Intercept and -3.18, - 3.50 and -4.15 for the one with the Intercept and the Trend. In consistence with the results for Albania, also in the case of Turkey, the residuals from all the regressions are stationary, thus, it is concluded that cointegration exists between each indicator of financial development and economic growth.

5.3. Granger Causality tests

Considering that both variables (finance and growth) move together in the long-run, we can proceed with the VECM which includes the error correction term (ECT). The presence of the ECT, will enable us to capture the long-run cointegration properties of the series (Brooks, 2008). Results for Albania are outlined in Table 5.

Tuble 2. VIIIV VECTOR Results for Anount										
	LGDP	DEPAL	LGDP	LOANS_D	LGDP	PRCRAL	LGDP	M2AL	LGDP	CARAL
	coefficient	coeff								
	[t-ratio]	[t-ratio]								
EG(-1)	0.05	-0.14	-0.49	0.38	-0.44	0.29	-2.45	4.83	-0.59	0.02
	[0.18]	[-0.21]	[-2.79]	[1.42]	[-2.5]	[1.1]	[-4.64]	[4.5]	[-4.51]	[2.85]
(-2)	-1.14	2.13	-0.56	0.51	-0.48	0.4	-0.55	0.83	-0.35	0
	[- 3.85]	[3.18]	[-3.59]	[2.13]	[-3.14]	[1.72]	[-0.87]	[0.65]	[-2.78]	[2.11]
FD(1)*	0.06	-0.18	-0.12**	-0.52	-0.15	0.5	0.93	1.7	-47.6**	0.92
	[0.38]	[-0.53]	[-1.02]	[-2.84]	[-1.27]	[2.93]	[3.68]	[3.31]	[-1.4]	[7.24]
(-2)	0.48	0.99	-0.36**	0.09	0.35	0.05	-0.15	0.03	-38.6**	-0.07
	[3.44]	[3.13]	[-3.52]	[0.63]	[3.56]	[0.38]	[-0.53]	[0.06]	[-1.31]	[-0.65]
ECT(-1	-0.34	0.76	-0.28	0.5	-0.37	0.63	-0.09	0.15	-0.1	-0.02
	[-4.41]	[4.19]	[-2.6]	[3]	[-2.96]	[3.38]	[-2.34]	[1.78]	[-2.06]	[-1.9]

Table 5. VAR/VECM Results for Albania

* DEPAL and M2AL are included in the VAR in their first differences, LOANS_DEPAL and PRCRAL in their second differences, whereas CARAL in its level. The optimal lag structure of the model is two, selected by minimizing the Akaike Information Criteria (AIC)

** It is noted that the coefficients of the loans to deposits ratio and the banks' capital to assets ratio are negative, that is the variables have a negative effect on economic growth. However, the causal relationship between the latter and these two finance variables will be revealed by the Granger causality tests.

In order to investigate whether there is a causality relationship between finance and growth Wald χ^2 statistics are applied for dependent variables as a group to determine short-term causality whereas t-statistics of error correction term will specify whether there is long-run causality. Moreover, joint Wald χ^2 statistics are reported in order to determine Granger endogeneity. Results for Albania are outlined in Table 6.

	t-stat f	for ECT	Wa	$ld \chi^2$	Joint Wald χ^2	
	FD-EG	EG-FD	FD-EG	EG-FD	FD-EG	EG-FD
DLGDPAL, DDEPAL	4.21***	4.19***	13.03***	10.33***	13.03***	10.33***
DLGDPAL, DDLOANS_DEPAL	2.6***	3***	13.03***	5.37*	13.03***	5.37*
DLGDPAL, DDPRCRAL	2.9***	3.38***	12.95***	6.37*	12.95***	6.37*
DLGDPAL, DM2AL	2.34**	1.78*	13.8***	20.36***	13.8***	20.36***
DLGDPAL, CARAL	2.06**	1.9*	2.12	9.21**	2.12	9.21**

Table 6. Granger Causality tests for Albania

*, ** and *** indicate significance at the 10%, 5% and 1% level

As observed in Table 6, financial development Granger causes economic growth in Albania both in the short and in the long-run since four out of five measures of financial development significantly cause real GDP in the short term whereas all five have a significant effect on the latter in the long-run. The bank Capital to Assets ratio (CAR) seems to have an impact on growth only in the long-run. On the other hand, economic growth Granger causes financial development both in the short and in the long-run since all the Wald χ^2 statistics are significant along with the t-statics of the error correction terms. To this extent, we conclude that bi-directional causality exists between finance and growth for Albania, both in the short and in the long-run. Our results are similar to those of Dushku (2010) who performed the same study for Albania. However, the author failed to find feedback relationship between finance and growth in the short-term.

In the case of Turkey, results from the VAR/VECM are outlined in Table 7. It should be noted that for the robustness of the model, certain diagnostic tests are performed for both the countries. The results from the tests are available upon request.

											-			
	LGDP	DEP	LGDP	LOAN	LGDP	PR_CR	LGDP	M2TR	LGDP	CAR	LGDP	MARK	LGDP	STOCK
	coeff													
	t-ratio													
EG	0.36	4.15	0.16	2.64	0.16	2.47	0.43	4.02	0.14	0	0.44	0	0.14	0
	[2.14]	[1.97]	[1.09]	[1.57]	[1.1]	[1.53]	[2.85]	[2.05]	[1.06]	[0.25]	[3.87]	[0.59]	[1.18]	[1.36]
FD	0.01	0.09	-0.3**	0.11	0.03	0.12	0	0.09	-1.1**	0.3	20	0.2	7.19	0.54
	[0.82]	[0.57]	[-2.7]	[0.77]	[2.69]	[0.83]	[0.7]	[0.65]	[-1.2]	[2.2]	[2.36]	[1.47]	[1.6]	[4.45]
ECT	-0.79	8.66	-0.63	4.82	-0.63	4.69	-0.86	9.42	-0.3	0	-0.97	0	3.23	0
	[-6.77]	[5.93]	[-5.06]	[3.48]	[-5.1]	[3.52]	[-6.95]	[5.92]	[-3.39]	[0.22]	[-8.4]	[3.59]	[-5.09]	[1.88]

Table 7. VAR/VECM Results for Turkey

*All the variables are included in the VAR in their first differences, except for the MARK_CAP and STOCKS_TRD which are stationary at their level. The optimal lag structure of the model is one, selected by minimizing the Akaike Information Criteria (AIC).

** It is noted that the coefficients of the loans to deposits ratio and the banks' capital to assets ratio are negative, that is the variables have a negative effect on economic

growth. However, the causal relationship between the latter and these two finance variables will be revealed by the Granger causality tests.

	o or orange		ŕ		Joint Wold u ²		
	t-stat for ECT		w	ald χ^2	Joint Wald χ^2		
	FD-EG	EG-FD	FD-EG	EG-FD	FD-EG	EG-FD	
DLGDPTR, DDEPOSITTR	6.77***	5.93***	0.67	3.9	0.67	3.9	
DLGDPTR, DLOANS_DEPTR	5.06***	3.48***	7.35***	2.48	7.35***	2.48	
DLGDPTR, DPRIV_CREDITTR	5.1***	3.52***	7.23***	2.36	7.23***	2.36	
DLGDPTR, DM2TR	6.95***	5.92***	0.5	2.22	0.5	2.22	
DLGDPTR, DCARTR	3.39***	0.22	1.48	0.06	1.48	0.06	
DLGDPTR, MARK_CAP	8.4***	3.59***	5.61**	0.35	5.61**	0.35	
DLGDPTR, STOCKS_TRD	5.09***	1.88*	5.06**	2.69	5.06**	2.69	

 Table 8. Granger Causality tests for Turkey

*, ** and *** indicate significance at the 10%, 5% and 1% level

Results outlined in Table 8 suggest bi-directional causality between financial development and economic growth in the long-run since all the t-statics of the error correction term are significant (except for the t-statistic of the ECT when causality runs from real GDP to CAR). In the short-term though, findings do not suggest the same. Firstly, it results that only the loans to deposits ratio, the bank credit to private sector, the market capitalization and the value of the stocks traded as a share of GDP cause economic growth in the short-term. The other indicators of finance seem not to have a significant impact on growth in the short term. On the other hand, both the financial market development indicators included in our model, matter for economic growth in the short and in the long-run. As far as the effect of growth on finance is concerned, the insignificance of the Wald χ^2 statistics in all the growth-finance models examined, suggests that economic growth does not cause financial development in the short-term. To conclude, it can be assumed that bi-directional causality exists between the two variables in the long-run whereas in the short-term their relationship cannot be clearly established. Our findings are similar to those of Demirhan et al (2011) who concluded as well in a bi-directional causality between finance and growth in Turkey, but both in the short and in the long-run.

Apart from the causality, our study pinpoints the negative effect of two measures of banking sector development in economic growth, more precisely capital to assets ratio and loans to deposits ratio. Even though at first, it may be assumed that a relatively high bank capital guarantees a healthy banking system, CAR ratio was found to have a significant negative impact on economic growth for both Albania and Turkey, more precisely, the higher the banks' capital, the slower the economy's growth. Indeed, Luci (2002) argues that in the developing countries, the increase of the minimal capital requirement can influence the stability of the banking system by avoiding the emergence of new banks, supposed as "risky", rather than by improving the actual capital situation of the existing banks. To this extent, the economy may be negatively affected as well. Furthermore, a higher bank capital means less sound investments

undertaken by the banks; hence, the economy's growth may be harmed. On the other hand, a high proportion of loans towards deposits were found to have a significant negative effect on growth, in both the countries. Indeed, Gualandri et al (2009) address the cause of the recent global financial crises to the decreased liquidity of banks and their attempts to give loans by borrowing additional funds rather than by counting on their customers' deposits. As it results also in the case of Albania and Turkey, these attempts have caused the financial system to suffer for liquidity and the economy to decline.

6. Conclusions

The finance-growth nexus remains one of the most discussed topics in the international academic literature. The continuous interest on the link is related to the inconclusive results achieved by the scholars through the years. Moreover, recent research is focused on those countries whose investigation could not be performed earlier since their financial systems and economies have only experienced growth over the last decade.

In our study, we investigated the causality between finance and growth in two SEE countries, namely Albania and Turkey, over the 1998-2012 period. Despite their standard classification as developing countries, Albania and Turkey differ in terms of economic progress since the latter is experiencing high rates of growth and represents the 15th country in the world with the largest GDP, whereas this is not the case for Albania. To this extent, our key motivation was to discover the link that exists between finance and growth among various stages of development of the emerging economies. In order to examine the causal relationship, five indicators of financial system development were used in the case of Albania, whereas two more indicators of financial markets' development were added when the investigation was performed for Turkey. Cointegration tests suggested that there is a long-run relationship between each indicator of finance and the real GDP. Therefore, we examined the causality relationship between the two variables through the Vector Error Correction Model. Findings suggested that bi-directional causality exists between finance and growth in Albania, both in the short and in the long-run. The same was observed for Turkey but only in the long-run, since in the short-term the relationship between the two variables could not be clearly established. Hence, it can be concluded that the finance-growth link is stronger in the less developed economies rather than in the developed ones.

It was noted in our results that certain indicators of banking sector development such as loans to deposits ratio and bank capital to assets ratio have a significant negative impact on growth. To this extent, our findings have implications in terms of the financial system regulation in similar economies. Considering that high loans to deposits ratios and high capital to assets ratios may cause the economy to decline, regulators should encourage banks to keep relatively low levels of capital in their balance sheets and in the same time become more vigilant on the potential liquidity risk that is associated to the banks' investing incentives.

References

Ang, James B (2008) A Survey of recent developments in the literature of finance and growth. *Journal of Economic Surveys*, 22 (3), p.536-576.

Atakan, T et al. (2010) The impact of financial sector development on the economic growth of emerging countries: An empirical study for E-7 countries. *Journal of Modern Accounting & Auditing*, 6(12), p. 50-62.

Beck, Th and Levine, R (2001) Stock Markets, Banks and Growth: Correlation or Causality. World Bank Policy Research Working Paper No. 2670; Carlson School of Management Working Paper, p. 1-24.

BoA (2012) Financial Stability Report 2012, p.1-120.

Brooks, C (2008) *Introductory Econometrics for Finance*, 2nd ed., Cambridge University Press, New York.

Calderon, C and Liu, L (2003) The direction of causality between financial development and economic growth. *Journal of Development Economics*, 72(1), p.321-334.

Demetriades, P and Andrianova, S (2004) Finance and Growth: What we know and what we need to know. In Goodhart (ed) *Money*, *Finance and Growth*, Routledge.

Demetriades, P and Law, S (2006) Finance, Institutions and Economic Development. *International Journal of Finance and Economics*, 11 (3), p.245-260.

Demirhan, E et al. (2011) The Direction of Causality between Financial Development and Economic Growth: Evidence from Turkey. *International Journal of Management*, 28 (1), p.3-19.

Disbudak, C (2010) Analyzing the Bank Credit-Economic Growth Nexus in Turkey. *European Journal of Economics, Finance & Administrative Sciences*, 23, p.34-48.

Dornbusch, R and Reynoso, A (1989) Financial Factors in Economic Development. *American Economic Development*, 79 (2), p.204-210.

Dushku, E (2010) Financial Development and Economic Growth: The Albanian Case. *Bank of Albania Discussion Paper, No.32.*

Gualandri, E et al. (2009) Financial Crisis and New Dimensions of Liquidity Risk: Rethinking Prudential Regulation and Supervision. *Journal of Money, Investment and Banking*, 8, p.25-42.

Hondroyiannis, G et al. (2005) Financial Markets and Economic Growth in Greece, 1986-199. *Journal of International Financial Markets, Institutions and Money*, 15(2), p.173-188.

Hosny, Amr S (2012) Financial development and economic growth in Egypt. *Macro-economics & Finance in Emerging Market Economies*, 5(1), p.62-77.

IMF (2013) Hopes, Realities and Risks. World Economic Outlook, p. 149-176.

Kar, Muhsin et al. (2008) Trade Liberalization, Financial Development and Economic Growth in The Long Term: The Case of Turkey. *South East European Journal of Economics & Business (1840118X)*, 3 (2), p.25-38.

Khodzhibaeva, M and Nikolaidou, E (2012) Empirical investigation of credit risk determinants: The case of Tajikistan banking industry. *Proceedings of the* 7^{th} *South East European Doctoral Student Conference*, p.125-133.

King, R.G and Levine, R (1993) Finance and Growth: Schumpeter might be right. *Quarterly Journal of Economics*, 108 (3), p.717-738.

Levine, R (1997) Financial Development and Economic Growth: Views and Agenda. *Journal of Economic Literature*, 35(2), p.688-726.

Levine, R (2004) Finance and Growth: Theory and Evidence. NBER Working Paper Brown No. W 10766 .University - Department of Economics; National Bureau of Economic Research (NBER).

Levine, R and Zervos, S (1998) Stock Markets, Banks and Economic Growth. *American Economic Review*, 88(3), p.537-558.

Jarvis, Christopher (2000) The rise and fall of Albania's pyramid schemes. *Finance & Development*, 37(1), p. 46-50.

NBRT (2012) Financial Stability Report 2012, p.1-146.

Patrick, H (1966) Financial Development and Economic Growth in Underdeveloped Countries. *Economic Development and Cultural Change*, 14(2), p.174-190.

Rioja, F and Valev, N (2004) Finance and the Sources of Growth at Various Stages of Economic Development. *Economic Enquiry*, 42(1), p.127-140.

Rosseau, P and Wachtel, P (1998) Financial Intermediation and Economic Performance: Historical Evidence from Five Industrialized Countries. *Journal of Money, Credit and Banking*, 30 (4), p.657-678. Saibu, O et al. (2011) Causal Relationships between Financial Development, Foreign Direct Investment and Economic Growth the Case of Nigeria. *International Journal of Business Administration*, 2(4), p.93-102.

Soytas, U and Kucukkaya, E (2011) Economic Growth and Financial Development in Turkey: New Evidence. *Applied Economics Letters*, 18(4-6), p.595-600.

Key properties of the enterprise-universities interorganizational networks and the capacity for introducing innovations

Krystian Gurba

Chair of Economics, Faculty of Management and Social Communication, Jagiellonian University in Krakow, Poland krystian.gurba@uj.edu.pl

Abstract. The paper reports ongoing research of the capacity for introducing innovations in networks of universities and enterprises. The subject of study are horizontal, interorganizational networks, formed by universities and enterprises perform joint R&D activity leading to introduction of innovations. The aim is to verify hypothesis that selected, key properties of the networks of entrepreneurs and universities enable growth of the network effectiveness in terms of the capacity for introducing innovations. Data on the activity of the Polish universities and enterprises in networks obtained in the study, will allow to examine the occurrence of correlation and regression between indicators of innovation introduction capacity and parameterized properties the network. Data will also enable observation of concentration of specific properties in relation to the given indicator through the use of the method of cluster analysis.

Keywords: *interorganizational networks, innovations, universities, enterprise, properties*

1 Introduction

Research focuses on the innovation introduction capacity. It takes under consideration interorganizational networks formed to carry out joint R&D activities as a step leading to the innovation introduction (Frascati Manual, 2002). The subject of analysis will be network properties, because of their importance for the effective management.

According to the most common definition, being expanded version of the classical Schumpeterian concept, innovation is the implementation of a new or significantly improved product (good or service or process, a new marketing method, or a new organizational method) in business practice of organizations (Oslo Manual, 2008). Introduction of each type of innovation is one of the main factors of competitiveness, which, based on innovation, has a more long-lasting and stable nature.

Networks are the most appropriate form of action a turbulent environment. Cooperation facilitates adaptation to changing conditions, while uncertainty is reduced by division of risk, acceleration of the innovation activity process and synergy that provides skills not available for a single organization (Kelly, 2001).

Creating a synthesis of definitions existing in the literature, interorganizational networks being subject of this paper can be defined as formal systems consisting of at least two organizations linked by complex, reciprocal and relatively permanent ties, formed to achieve a certain objective (Sydow, 1998; Podolny, Page, 1998).

Direct cooperation with the private sector is considered as a task of modern, entrepreneurial universities. Therefore, an important research task is to determine the ability to manage the networks by enterprises and academia, and to create and configure connections that will increase network effectiveness.

It is defined as making decisions appropriate for the organization and their successful implementation (Griffin, 1996). Effectiveness can be measured as a degree to which the organization achieves its objectives (Stoner, Freeman, Gilbert, 1995). Effectiveness in the context of this article will therefore represent the capacity to achieve the aim of the network: introduction of innovations.

Capacity to introduce innovations means the ability to create and implement innovations. The most important European research measuring R&D activities and innovativeness is the Innovation Union Scoreboard (IUS). Consisting of 24 indictors it provides comparable results from each European country and allows for benchmarking of national economies (Innovation Union Scoreboard, 2013). In research, indicators based on IUS will be used to determine the capacity to introduce innovations, due to their comprehensive nature and the ability to adapt them to the testing of individual networks. Indicators, such as: expenditure on R&D, the number of employees engaged in R&D, the number of patent applications, the number of joint scientific publications and license revenue, will be used to examine the capacity on network-level and on the level of individual organizations.

In the course of the study full set of properties of networks will be identified. Examples of network properties that compose it are: the size, the period of performance, the budget on R&D activities, the strength of the links in the network, the integrity, the density, the size of the organizations within the network, their experience in R&D cooperation, the decision-making process. The study aims to determine which properties are important for the effective management of networks in the context of capacity to introduce innovations.

2 Literature Review

Main theoretical concepts on innovation, being basis for planned research are Open Innovation model (Chesbrough, 2003), concept of National Innovation Systems (Freeman, 1991) and Triple Helix theory (Etzkovitz, Leyderdorff, 1998). They lead to following questions: which endogenous factors affect the capacity to cooperate effectively to introduce innovations, how to achieve synergy effect in network and what is the role of the government and intermediate organizations in innovation introduction.

The issue of network and networking in activity of organization has been and is being developed in management sciences primarily through the researchers representing strategic management approach. Industrial network theory group (Haakanson, Snehota) have concentrated their research on organizations with blurred boundaries and the reasons and methods of formation of the network. Strategic network group (Jarillo, Gulati, Nohria, Zaheer), sees them as a deliberate arrangements, initiated and coordinated by the organization aimed at obtaining a competitive advantage (Tikkanen, Halinen, 2003). The ties between organizations and network creation are subject of study also in institutional economics.

Query revealed lack of studies on the effectiveness of networks in the context of innovation introduction. Fritsch and Kauffeld-Monz provided results of research conducted among nearly 300 German companies and research organizations active in regional innovation networks (Fritsch, Kauffeld-Monz, 2010). Researchers examined them in terms of ties between network partners and the transfer of knowledge, but the network properties influence on innovation introduction has not been tested.

On the other hand, studies like analysis of the Polish university-enterprise network (Olechnicka, 2012), focused on the role of their geographical location and the impact on the region in which they operate. Less explored issue is the significance of endogenous factors relating to the characteristics the network itself.

Selected interorganizational network properties, concerning relations in the network, were subject of research by Sydow (Sydow, 1998). Powell discusses diversity of organizations in network and centrality of the network (Powell, 1998). Furthermore, the strength of ties and its impact on access to information, as well as coherence and scope of the network, were a subject of studies (Reagan, McEvily, 2003; Granovetter, 1982). Aim of the research is to verify if conclusions of above-mentioned studies are applicable to innovation introduction capability in networks.

3 Proposed Methodology

The research applies various analysis methods and research techniques.

The main qualitative method used in the study is query, systematization and comparative analysis of existent knowledge: literature, legal acts and other documents such as financing programmes, guidelines for network creation, consortium agreements and internal rules of organizations forming the network. Additionally, as an introduction to main research, interviews with representatives of key universities cooperating with entrepreneurs may be conducted,

Quantitative research will be carried out using available statistical data on the networks and data collected in the survey. They will enable to investigate the changes in indicators of capacity for introducing innovation in networks and their relation with the properties of these networks.

Quantitative research will consist of the following phases:

- 1. Collecting data about population of universities-enterprise networks in Poland from public research funding organizations.
- 2. Disigning the size of research sample and selection of the representative group.

- Preparation and dissemination of the questionnaire (Computer Assisted Web Interview – CAWI), with closed questions about network properties (those of a nonnumeric character will be parametrized using the 5-level Likert scale).
- 4. Discussion of the results, followed by verification of the main research hypothesis.

Selected network properties will be examined through inquiries, on full population of Polish enterprise-universities networks. Study of other properties will have a character of partial, representative research. The samples shall consist of networks with participation of all defined types of universities and shall represent proportionally different technology sectors. The results of those samples will be generalized to the entire population through statistical inference.

Statistical analysis will be used for measure of correlation and multidimensional regression. Parameterized network properties will constitute independent variables and the individual performance indicators – the dependent variables. The analysis will cover all pairs of variables: network property-performance indicator. To remove the effects of other variables, partial correlation analysis shall be used.

Obtained data will also be used to perform cluster analysis – a tool for exploratory data analysis (Tan, Steinbach, Kumar, 2006). It will enable identification of the groups of properties that are most important for obtaining high levels of the network capacity for innovation and a spatial concentration of network properties connected with each indicator of innovation introduction capacity.

Additionally the social network analysis (a research tool allowing to map and analyse relations between members of the group) will be used. to visualize the links between organizations being part of the networks in Poland.

The research methods used in study will result in achieving following effects:

- to identify and organize the complete catalog of unique network properties;
- to measure the effectiveness with the capacity to introduce innovations indicators;
- to develop a model of interorganizational network research in the context of network properties for use in further research;
- to determine the importance of each network property for the effectiveness network and benchmark the properties.

4 Conclusions

Proposing an instrument which facilitates achievement of the objectives of the network, has a considerable practical importance. Innovative nature of the research will rely primarily on linking the properties of the network with the level of capacity for the introduction of innovations, as well as the application of new methods of statistical analysis: cluster analysis and social network analysis. Moreover, a model of interorganizational network management, being a product of the research will be useful for designing networks and increasing their effectiveness, as well as for future studies.

References

Chesbrough H. W. (2003) The era of open innovation, *MIT Sloan Management Review*, Vol. 44 (3).

Etzkowitz H., Leydesdorff L. (1998) The Triple Helix as a Model for Innovation Studies, *Science & Public Policy*, Vol. 25(3).

(2002) Frascati Manual. The Proposed Standard Practice for Surveys of Research and Experimental Development, OECD.

Freeman C. (1991) Network of innovators: a synthesis of research issues, *Research Policy*, Vol. 20, p. 499-514.

Fritsch M., Kauffeld-Monz M. (2010) The impact of network structure on knowledge transfer: an application of social network analysis in the context of regional innovation networks, *The Annals of Regional Science*, Vol. 44, Issue 1, p. 21-38.

Granovetter M.S. (1982) The strength of weak ties: a network theory revisited, in Marsden P.V., Lin N., (eds), *Social Structure an Network Analysis*, Sage, Beverly Hills,

Griffin R. W. (2002) Fundamentals of Management – Core Concepts and Applications, 3rd Ed., Houghton Mifflin Company.

Kelly K. (1998) New Rules for the New Economy: 10 Radical strategies for a Connected World, Penguin, Harmondsworth.

Olechnicka A. (2012) Potencjał nauki a innowacyjność regionów, Wydawnictwo Naukowe Scholar, Warszawa.

(2008) Oslo Manual - proposed guidelines for collecting and interpreting technological innovation data, OECD.

Podolny J.M., Page K.L. (1998) Network forms of organization, *Annual Review of Sociology*, no. 24, p. 57-76.

Powell W.W. (1998) Learning from collaboration: knowledge and networks in the biotechnology an pharmaceutical industries, *California Management Review*, vol. 40 (3).

Reagans R., McEvily B. (2003) Network structure an knowledge transfer: the effects of cohesion and range, *Administrative Science Quarterly*, 48(2), p. 240-267.

Stoner J.A., Freeman R.E., Gilbert D.R. (1995) *Management*, Prentice Hall International Limited, London.

Sydow J. (1998) Understanding the constitution of interorganizational trust, in Lane Ch., Bachman R. (eds) *Trust Within and Between Organizations*, Oxford University Press, Oxford.

Tan P.-N., Steinbach M., Kumar V. (2006) Introduction to Data Mining, Addison-Wesley.

Tikkanen J., Halinen A. (2003) Network approach to strategic management – exploration to the emerging perspective, 19th Annual IMP Conference, Lugano.

(2013) *Innovation Union Scoreboard*, http://ec.europa.eu/enterprise/policies/innovation/files/ius-2013_en.pdf [12.04.2013].

The Role of Networking on SMEs Innovation: the case of Kosovo

Flaka Ismaili1

¹South-East European Research Centre (SEERC), University of Sheffield flismaili@seerc.org

Abstract. This short paper aims to explain the rationale behind the intended research and propose a suitable methodology to conduct the study. The issue developed in the paper is to see whether there is some kind of a relationship between external networking and innovation, in SMEs located in Kosovo. The study will be able to fill in the literature gaps concerning the role of networking on SMEs on the context of developing countries; it will provide a study using ego-network approach to explain the relationship between network structure and innovation. The proposed methodology is a mixed methods approach, in order to get a more in-depth view of the issue.

Keywords: *external networking, network structure, SME innovation, Kosovo*

1. Introduction

In recent years, policy makers have shown growing interest in promoting small and medium enterprises (SMEs in further text) since they are seen as key drivers of economic growth and employment (Rothwell, 1991; Tambunan, 2011). And the primary mean through which they are expected to achieve this is by developing and encouraging innovation (Radas & Božić, 2009; OECD, 2004).

The importance of SMEs in creating economic growth and employment is also recognized as essential for both developed and developing economies, thus pushing them into finding different ways to encourage SME innovation (Radas & Božić, 2009). A lot of research has been conducted the last decades either to identify the ways that SMEs can be helped to innovate or to identify the barriers that hinder innovation (Keizer, Dijstra, & Halman, 2002; Madrid-Guijarro, Garcia, & Auken, 2009; Rivera-Vazquez, Ortiz-Fournier, & Flores, 2009; Radas & Božić, 2009; Zhu, Wittmann, & Peng, 2012). And some of the studies have pointed out that networking plays a major role on the firm's innovative performance (Hoang & Antoncic, 2003). It is believed that networking can offer SMEs the opportunities that their internal resources are not able to (Chesbrough H. W., 2007; Tomlinson & Fai, 2011).

As it was put above, SMEs are considered as very important to economical processes, and as such their success is deemed necessary. Accordingly, this paper will focus on arguing that networking plays a vital role on the success of SMEs. Additionally, it will explore the relationship between different forms of innovations (taking into account the Oslo Manual definition of innovation) and external networks.

The difference of this paper in relation to others is that it suggests a study on a less developed country (Kosovo), where the economy is in transition and the country is a post-conflict region. The topic of networking, in relation to SME innovation, is relatively explored; however, there is a gap on less developed countries and in postconflict regions the research is almost nonexistent. Additionally, the relationship between different forms of innovation and networking is understudied, especially though ego-network approaches.

Nevertheless, this paper will consist of several parts which will present the topic. The first part will be the literature review which will argue and critically evaluate the relevance and relationship of the above laid themes. Then the research questions and aims of the proposed study will be laid out. After that a methodology will be proposed, and the paper will end with conclusions.

2. Literature Review

Innovation is viewed as both a process and an output (Gronum, Verreynne, & Kastelle, 2012; Salavou & Lioukas, 2003). According to Love, Roper, and Du (2009), innovation is considered as an output when it is the result of the innovation process and the implementation of types or forms of innovations, such as introduction of a new product, process, service or method. Whereas, the process of innovation refers to the chain of events where people interact with one another to develop and implement their new and innovative ideas (Gronum, Verreynne, & Kastelle, 2012). It is believed that both the innovation process and its resulting outputs can affect the firm's performance and success (Rosenbusch, Brinckmann, & Bausch, 2011).

In order to understand and measure the relationship between innovation and firm's performance or any other variable, many scholars have used the definition provided by the Oslo Manual, (OECD, 2005), which is as follows:

"Innovation is the implementation of any new or significantly improved product (goods or services), operational processes (methods of production and service delivery), any new marketing methods (packaging, sales and distribution methods), or new organizational or managerial methods or processes in business practices, workplace organization or external relations" (p.46).

Furthermore, encouraging innovation in SMEs is believed to be a very important issue for the economy; thus a lot of studies have been carried out in order to find the factors that best contribute to the innovative attempts by SMEs (Radas & Božić, 2009; Keizer, Dijstra, & Halman, 2002). According to Keizer et al. (2002), the factors that can impact the innovation of a firm, can be divided into internal and external factors, where internal factors refer to the policies and characteristics of an SME, and external factors refer to the opportunities that SMEs can draw upon from its environment.

Some of the internal factors, identified through decades of research that affect the innovation in SMEs, are: the strategy, different kinds of policies, management structures, investments in R&D and marketing (Radas & Božić, 2009). However, regarding external factors, they are considered as: collaboration and cooperation with other agents (networking) and exploiting financial resources (Keizer, Dijstra, & Halman, 2002).

Over the last decades, there has been an organized change in the way firms carry out innovatory activities (Zeng, Xie, & Tam, 2010). Particularly, there has been a

remarkable rise of the usage of external networks by enterprises of all sizes (Hagedoorn, 2002). It is widely recognized that the critical resources of a firm may be expanded beyond the traditional boundaries and that firms tend to extract the resources and expertise from others (Tomlinson & Fai, 2011). This extraction of others' resources and expertise is usually done through the firm's business networks that commonly include its suppliers, main customers, outlets and in some cases cooperation with competitors or alternatively known as co-opetition (p.319). Internal and external networking are nowadays considered as key strategies because they are believed to facilitate and speed up the flow of information, resources and trust necessary for innovation (Dewick & Miozzo, 2004). Additionally, Diez (2002) suggested that since SMEs do not have many resources or R&Ds, and are faced more with uncertainty and barriers to innovation, networks can balance the rising insecurity, that come from using new technologies or from development, and can reduce the uncertainty in innovation. Moreover, this era is considered as the era of "open innovation", according to Chesbrough (2003), where the author also states that the firms largely rely on external sources for creating successful innovation.

Different studies and authors consider two categories of networking: *internal networking* and *external networking* (Helble & Chong, 2004; Gulati, Nohria, & Zaheer, 2000). External networking refers to the relationships formed with external parties, like other firms, government institutions, competitors, different agencies, universities etc, and internal networking refers to the cooperation and relationship between internal agents within a firm (Helble & Chong, 2004). Vereecke et al. (2002) identified three types of relationships in internal networking. The first one was the relationship of human resources and its flow between different sectors. The second one was the relationship of innovation, while focusing on how much the sources are shared. And the third one was the relationship of information, focusing on the degree of freedom and information flow. Moreover, concerning the relationships on external networking, a similar distinction was applied by many authors, such as: Helble & Chong (2004), or Andersson et al. (2002).

In the literature concerning networks, three elements of networks were identified: 1) the nature of content exchanged between parties; 2) the governance mechanisms; and 3) and the network structure, (Hoang & Antoncic, 2003). A lot of researchers and scholars have used these three components as key elements in the models that try to explain the impact of networks in entrepreneurial outcomes, or to explain the process of network development during innovation attempts (Hoang & Antoncic, 2003; Hite, 2003; Amit & Zott, 2001).

The first model that characterizes the features of the networking by its content focuses on the exchanged resources between actors or parties. According to Hoang & Antoncic (2003), the networking based on the nature of its content, is basically measured by the amount of multiple resources exchanged.

The second model explains networking through mechanisms supporting exchange in order to coordinate and manage relationship, especially by trust (Hoang & Antoncic, 2003). Others have explained network governance by highlighting the importance of relying on open ended contracts, or maintaining reputation (Provan & Kenis, 2008; Jones, Hesterly, & Borgatti, 1997).

And the third model is the network structure, which is defined as the direct and indirect ties between actors or as the relationship between actors both interpersonal and inter-organizational (Hoang & Antoncic, 2003). Some of the measures used for network structures are size, centrality, density, and strong/weak bridging ties or relationship strength (Wasserman & Faust, 1994; Hoang & Antoncic, 2003; Jones & Volpe, 2011)

A number of researchers identified *size* as the most intuitive measure for networking structure, where they define it as the extent to which resources can be accessed at the individual level (the entrepreneur) or organizational level (Hansen, 1995; Baum, Calabrese, & Silverman, 2000). Another measure was the *centrality*, which refers to the ability to access and control resources though direct and indirect links (Hoang & Antoncic, 2003). However, scholars like Jones & Volpe (2011) consider the network size and network centrality as the same where throughout their papers, they use the terms interchangeably. Furthermore, network *density* refers to the extent actors in a network know one another and how interconnected their network is (Jones & Volpe, 2011). It is considered as a very important measure since it facilitates communication, trust and support among actors (Feld & Carter, 1998; Jones & Volpe, 2011). And the last but not least, *relationship strength*, is yet another measure that is defined as the frequency of interaction, duration and closeness, which is more or less the emotional intensity of a tie between actors or parties (Jones & Volpe, 2011).

Nevertheless, there are a handful of studies concerning networking and innovation in SMEs. According to different authors, most of the studies are focused on vertical supply chain networks, or they try to identify different measures of innovation, and the nature and level of network ties (Tomlinson & Fai, 2011; Huizingh, 2011). Similar studies have found a positive relationship between enterprises that had cooperation with both client firms and suppliers over innovation (De Propris, 2002; Freel & Harrison, 2006). However, De Propris (2002) also found that concerning process innovation, only supplier co-operation was noticeable. Additionally, a study on 1300 Spanish SMEs identified the vertical technological collaboration as the most important factor for improving firms' innovativeness (Nieto & Santamaria, 2010), while another one on 500 SMEs of six different European countries found that cooperation with both buyers and suppliers, concerning the design, positively helped innovation (Lasagni, 2012).

Some researchers also found that network cooperation involving external actors or sources, such as universities, research institutes, suppliers, customers etc, had a positive and significant affect on the process of creating knowledge and innovation (Brioschi, Brioschi, & Cainelli, 2001; Nieto & Santamaria, 2007; Bullinger, Auernhammer, & Gomeringer, 2004). Similarly, Nieto & Santamaria (2007), in their longitudinal study of Spanish manufacturing SMEs found that the greatest impact on the degree or level of innovation came when multiple collaborative partners where used for networking.

On the other hand, there are only few studies from developing countries focusing on the relationship between networking and SMEs' innovativeness. Hadjimanolis (1999), on a study of SMEs in Cyprus, then a developing country, identified the types of inter-organizational collaborations as being vertical and horizontal (having strong relationships with customers, suppliers and different associations), and he also identified factors that interfere with innovation. Furthermore, a survey carried out in China by Liefner et al. (2006), found that SMEs that cooperated with foreign companies, got help for creating new ideas and enter markets, and collaboration with universities was used only for designing new products. Yet another study carried out in Africa indicated that those SMEs that were collaborating with each other had more activities accompanied by innovation (Biggs & Shah, 2006).

Nonetheless, the studies conducted in Kosovo mostly argue and present the importance of innovation both as a process itself and as part of SME, or try to identify factors that play a role on its process (Roper, 2009; Harms, Wdowiak, & Schwarz, 2010; Soini & Veseli, 2011). For example one of those studies identified that institutional related barriers had the most negative impact on SMEs' activities, such as: informal activities by inappropriate professionals, questionable ethics of officials and the poor enforcement of regulations (Peci, Kutllovci, Tmava, & Shala, 2012).

It can be concluded that the above literature offers an ample review on innovation and networking in SMEs. However, it can be clearly seen that there is lack of research on the role of networking on SMEs on the context of developing countries. Besides this, there are no studies that clearly identify the roles of each networking agency has on different innovation forms and nonetheless in developing countries, or Kosovo for that matter. Therefore, it is under this perspective that this paper proposes to undertake such a study.

3. Research questions and aims

As also mentioned at the beginning the aim of this paper is to argue the importance of exploring the potential relationship external networking can have on different forms of SME innovation. Another scope that it is proposed to be explored is the context of developing countries, where it proposes that the study be conducted in Kosovo, because of the total lack of studies about the relationship between networking and innovation.

In light of the above literature on previous inquiries and the offered definitions, our paper regards the following research questions to be considered for the study:

- 1) Identify the types of external networks used by SMEs in Kosovo?
- 2) Examine the relationship between external networks and the forms of innovation¹?
- 3) Investigate the level of impact different network structures (size/centrality and relationship strength) have on innovation.

4. Proposed Methodology

It is proposed that the study adapt a mixed-method approach, using both quantitative and qualitative research. The sample can consist of SME chosen for their particular nature as innovative. It is suggested that the target population include 2 groups, employees for questionnaires and managers or owners for interviews. An SME should be considered a firm that has 10 - 250 employees, as also recommended in other studies (Tomlinson & Fai, 2011). Experienced researches suggest that a combined methodology, employing both quantitative and qualitative approaches, can offer a clearer and in depth picture of the issue at hand (Tashakkori, 2003). Moreover, mixed-

¹ The forms of innovation are strictly considered as in the definition of innovation by the Oslo Manual, as also presented on the Literature Review section.

methods are usually used to increase the comprehension and understanding of an issue through corroborating conclusions, or by starting and facilitating new ways of thinking about a particular theme (Bazeley, 2002).

Sample and target population

Going back to the sample and target population, the region where the research should take place is Kosovo. Since Kosovo is a relatively small country with a population of almost 2 million, for a mixed-method approach 6 to 10 SMEs will be more than enough. It is recommended that SMEs be identified and grouped based on their sector and their history of innovativeness.

In Kosovo, according to the Ministry of Trade and Industry, SMEs are mostly concentrated in: retail (50%); transport, storage and distribution (14%); food, beverages and tobacco (9%); and, hotels and restaurants (9%) (Ministry of Trade and Industry, 2012). Thus, it is suggested that retail and wholesale sector is a good specific industry for the study, since it comprises the largest percentage of developed sector of SMEs.

Measurements

However, regarding measurements, there are a handful of studies that can be used as reference. For example, concerning measuring *innovation*, the Oslo Manual (2005) guidelines and suggestion can be used. Additionally, measuring the extent firms use *networks* can be done following the example of Zeng, Xie, & Tam (2010), where they asked the respondents to indicate the level their firm cooperates with different partners (suppliers, government institutions, university or research institutions, customers, etc), while using Likert scale, ranging from 1 to 5.

Nonetheless, to measure the components of *network structure* and their impact, the method used by Jones & Volpe (2011) can be adopted. To measure *network size or centrality*, they used Ibarra's (1993) method, but instead of using 5 units, they used only three, which were: 1. "with whom you discuss what is going on in the organization", 2. "who are important sources of professional advice, whom you approach if you have a work-related problem or when you want advice on a decision you have to make", 3. "people who have contributed most significantly to your development". Though, in order to measure *relationship strength*, they used a one item Likert – type scale, were the participants were asked to "indicate the closeness of their relationship with them" by placing an 'X' in one of the boxes marked Very Close, Close, Not Very Close, Distant.

On the other hand, since the research questions try to explore relationships, it is suggested to follow the examples of above mentioned authors and their studies. Jones & Volpe (2011) did the following tests: means, standard deviations, correlations (for a clearer profile of the sample and target population), hierarchical multiple regression (for determining relationship) and multicollinearity. Similarly, Zeng, Xie, & Tam (2010) used means, (M), standard deviations (S.D.), internal consistencies Cronbach's α), and since they had latent variable they used factor analysis (different from the previous mentioned authors). In this study variables are observable and there are no hidden ones, so factor analysis is not needed.

Limitations

It should also be kept in mind that to this study, certain limitations can arise. First the researcher needs to be careful for language barriers. Second in a country with a post conflict history and a transitioning economy it will be hard to identify the SMEs in question, since earlier data might be lost and there are very few institutions that keep track on the prospect of SMEs. Another limitation might be the time constrain. A cross sectional study will only capture measures of different samples at one point in time, and the understanding of changes can be limited.

5. Conclusions

The literature and different studies suggest that there is enough evidence to believe that networking can help SMEs in the innovation process. As elaborated above sufficient studies have resulted in showing that there was a positive relation between networking, and innovative success. Some authors claimed that networks can help SMEs innovate because they can offer resources that they lack. Some studies found that using external sources lead to successful innovations, and some others found that strong ties or relationships with other firms or parties, can lead to more and appropriate resources. Other authors also identified factors that affect innovation on SMEs, and some others identified the barriers they have. Among the positive effecting factor was networking also.

Empirical research was offered that showed studies, in both developing and developed countries, regarding networking and innovation in SMEs. Most of the studies on developed countries were focused on identifying different measures of innovation, and the nature and level of network ties. Nonetheless, regarding developing countries there were only a few and they mostly focused on identifying factors that affect innovation. Though, there were also some that found that cooperation with foreign firms helped SMEs innovate (a study in China), and that when SMEs collaborate with each other tend to have more innovative activities (a study in Africa).

Subsequently, it was seen that there are some gaps in the literature concerning studies employing ego-network approach involving network size, density and strength of ties in regards to innovation, and studies conducted in transitioning economy and developing countries.

Therefore in conclusion, exploring whether there is a specific relationship between external networking and innovation, in a country like Kosovo, will significantly contribute to the academic knowledge and practice.

References

Amit, R., & Zott, C. (2001). Value creation in E-business. *Strategic Management Journal*, 22, p. 493–520.

Andersson, U., Forsgren, M., & Holm, A. (2002). The strategic impact of external networks: subsidiary performance and competence development in the multinational corporation. *Strategic Management Journal*, 23(11), p. 979– 996.

- Baum, J. A., Calabrese, T., & Silverman, B. S. (2000). Don't go it alone: alliance network composition and startups' performance in Canadian biotechnology. *Strategic Management Journal*, 21, p. 267–294.
- Bazeley, P. (2002). Issues in Mixing Qualitative and Quantitative Approaches to Research. 1st International Conference - Qualitative Research in Marketing and Management (p. 141-156). UK, Palgrave Macmillan.
- Biggs, T., & Shah, M. K. (2006). African SMEs, networks, and manufacturing performance. *Journal of Banking and Finance*, *30*(11), p. 3043–3066.
- Brioschi, F., Brioschi, M. S., & Cainelli, G. (2001). From the industrial district to the district group: an insight into the evolution of local capitalism in Italy. *Regional Studies*, 36(9), p. 1037–1052.
- Bullinger, H. J., Auernhammer, K., & Gomeringer, A. (2004). Managing innovation networks in the knowledge-driven economy. *International Journal of Production Research*, 42(17), p. 3337–3353.
- Chesbrough, H. (2003). *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Boston: Harvard Business School Press.
- Chesbrough, H. W. (2007). Why companies should have open business models. *MIT* Sloan Management Review, 48, p. 22-28.
- De Propris, L. (2002). Types of innovation and inter-firm co-operation. *Entrepreneurship and Regional Development*, 14(4), p. 337–353.
- Dewick, P., & Miozzo, M. (2004). Networks and innovation: sustainable technologies in Scottish social housing. *R & D Management*, *34*(4), p.323–333.
- Diez, J. D. (2002). Metropolitan innovation systems: a comparison between Barcelona, Stockholm, and Vienna. *International Regional Science Review*, 25(1), p. 63–85.
- Fariselli, P., Oughton, C., Picory, C., & Sugden, R. (1999). Electronic Commerce and the Future for SMEs in a Global Market-Place: Networking and Public Policies. *Small Business Economics*, 12(3), p. 261–275.
- Feld, S. L., & Carter, W. C. (1998). Foci of activity as changing contexts for friendship. In R. G. Adams, & G. Allan, *Placing friendship in context* (p. 136–152). Cambridge, Cambridge University Press.
- Freel, M. S., & Harrison, R. T. (2006). Innovation and cooperation in the small firm sector: evidence from Northern Britain. *Regional Studies*, 40(4), p. 289–305.
- Gronum, S., Verreynne, M.-L., & Kastelle, T. (2012). The Role of Networks in Small and Medium-Sized Enterprise Innovation and Firm Performance. *Journal of Small Business Management*, 50(2), p.257–282.
- Gulati, R., Nohria, N., & Zaheer, A. (2000). Strategic networks. *Strategic Management Journal*, 21(3), p. 203–215.
- Hadjimanolis, A. (1999). Barriers to innovation for SMEs in a small less developed country (Cyprus). *Technovation*, 19(9), p. 561–570.
- Hagedoorn, J. (2002). Inter-firm R&D partnerships—an overview of patterns and trends since 1960. *Research Policy*, *31*(4), p. 477–492.
- Hansen, E. L. (1995). Entrepreneurial networks and new organization growth. *Entrepreneurship: Theory and Practice, 19*(4), p. 7-19.
- Harms, R., Wdowiak, M. A., & Schwarz, E. J. (2010). The role of entrepreneurshipand technology transfer-oriented initiatives in improving the innovation

system in the Western Balkans. International Journal of Business and Globalisation, 4(1), p. 18-34.

- Helble, Y., & Chong, L. C. (2004). The importance of internal and external R&D network linkages for R&D organisations: evidence from Singapore. *R&D Management*, 34(5), p. 605-612.
- Hite, J. M. (2003). Patterns of multidimensionality among embedded network ties: a typology of relational embeddedness in emerging entrepreneurial firms. *Strategic Organization*, 1(1), p. 9-49.
- Hoang, H., & Antoncic, B. (2003). Network-based research in entrepreneurship: a critical review. *Journal of Business Venturing*, 18, p. 165–187.
- Huizingh, E. K. (2011). Open innovation: state of the art and future perspectives. *Technovation*, 31(1), p. 2–9.
- Jones, C., & Volpe, E. H. (2011). Organizational identification: Extending our understanding of social identities through social networks. *Journal of Organizational Behavior*, 32, p. 413-434.
- Jones, C., Hesterly, W. S., & Borgatti, S. P. (1997). A general theory of network governance: exchange conditions and social mechanisms. *Acad. Manage. Review*, 22(4), p. 911–945.
- Kaufmann, A., & Todtling, F. (2000). Science–industry interaction in the process of innovation: the importance of boundary crossing between systems. *Research Policy*, *1241*, p. 1-14.
- Keizer, J., Dijstra, L., & Halman, J. I. (2002). Explaining innovative efforts of SMEs. An exploratory survey among SMEs in the mechanical and electrical engineering sector in The Netherlands. *Technovation*, 22, p. 1–13.
- Lasagni, A. (2012). How can external relationships enhance innovation in SMEs? New evidence for Europe. *Journal of Small Business Management*, 50(2), p. 310–339.
- Liefner, I., Hennemann, S., & Xin, L. (2006). Cooperation in the innovation process in developing countries: empirical evidence from Zhongguancun, Beijing. *Environment and Planning*, 38(1), p. 111–130.
- Love, J. H., Roper, S., & Du, J. (2009). Innovation, Ownership and Profitability. International Journal of Industrial Organization, 27, p. 424–434.
- Madrid-Guijarro, A., Garcia, D., & Auken, H. V. (2009). Barriers to Innovation among Spanish Manufacturing SMEs. *Journal of Small Business Management*, 47(4), p. 465–488.
- Ministry of Trade and Industry. (2012) *SME Development Strategy for Kosovo*. Ministry of Trade and Industry, Prishtina.
- Nieto, M. J., & Santamaria, L. (2007). The importance of diverse collaborative networks for the novelty of product innovation. *Technovation*, 27(6–7), p. 367–377.
- Nieto, M., & Santamaria, L. (2010). Technological collaboration: bridging the innovation gap between small and large firms. *Journal of Small Business Management*, 48(1), p. 44–69.
- OECD. (2004). *Global Knowledge Flows and Economic Development*. Paris: OECD Publications.

- OECD. (2005). Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data (3rd ed.). Paris: Joint Publication of the OECD and the Statistical Office of the European Communities.
- Peci, F., Kutllovci, E., Tmava, Q., & Shala, V. (2012). Small and Medium Enterprises Facing Institutional Barriers in Kosovo. *International Journal of Marketing Studies*, 4(1), p. 95-107.
- Provan, K. G., & Kenis, P. (2008). Modes of network governance: Structure, management, and effectiveness. *Journal of public administration research* and theory, 18(2), p. 229-252.
- Radas, S., & Božić, L. (2009). The antecedents of SME innovativeness in an emerging transition economy. *Technovation*, 29(6), p. 438 450.
- Rivera-Vazquez, J. C., Ortiz-Fournier, L. V., & Flores, F. R. (2009). Overcoming cultural barriers for innovation and knowledge sharing. *Journal of Knowledge Management*, 13(5), p. 257-270.
- Roper, S. (2009). Innovation in Transition: A Comparison of the Innovation Potential of Incumbent Firms and Innovative Start-ups in the Western Balkans. Warwick, Centre for Small and Medium Sized Enterprises, Warwick Business School.
- Rosenbusch, N., Brinckmann, J., & Bausch, A. (2011). Is Innovation Always Beneficial? A Meta-Analysis of the Relationship between Innovation and Performance in SMEs. *Journal of Business Venturing*, 26(4), p. 441–457.
- Rothwell, R. (1991). External networking and innovation in small and medium-sized manufacturing firms in Europe. *Technovation*, 11(2), p. 93-112.
- Salavou, H., & Lioukas, S. (2003). Radical Product Innovations in SMEs: The Dominance of Entrepreneurial Orientation. *Creativity and Innovation Management*, 12(2), p. 94–108.
- Soini, E., & Veseli, L. (2011). Factors influencing SMEs growth in Kosovo. Unpublished Bachelor's Thesis. Turku University of Applied Sciences, Helsinki.
- Tambunan, T. T. (2011). Development of small and medium enterprises in a developing country: The Indonesian case. *Journal of Enterprising Communities*, 5(1), p. 68-82.
- Tashakkori, A. T. (2003). A handbook of mixed methods in the social and behavioural science. Thousand Oaks: Sage.
- Tomlinson, P. R., & Fai, F. M. (2011). The nature of SME co-operation and innovation: A multi-scalar and multi-dimensional analysis. *International Journal of Production Economics*, 141(1), p. 316–326.
- Wasserman, S., & Faust, K. (1994). *Social Network Analysis*. New York, Cambridge University Press.
- Zeng, S. X., Xie, X. M., & Tam, C. M. (2010). Relationship between cooperation networks and innovation performance of SMEs. *Technovation*, 30(3), p.181-194.
- Zhu, Y., Wittmann, X., & Peng, M. W. (2012). Institution-based barriers to innovation in SMEs in China. Asia Pacific Journal of Management, 29(4), p. 1131-1142.

Examination of Competition Policy Effects on the Retail Market Development in Transition Countries by Applying the Panel Model

Siniša Milošević¹

¹Faculty of Economics, University of Belgrade, Serbia sinisa.milosevic3@gmail.com

Abstract. The transit process of countries in transition from planned to market oriented economies among other things includes the introduction of a regulatory body and regulatory frame to support competition policy. Therefore, competition policy has become an integral part of the economic policy in most countries in transition. In literature, there are numerous researches that examine the effect of competition policy on the economy as a whole. In spite of that, the subject of this research will be the examination of effects of competition policy to the development of retail industry. The hypothesis from which we began is that implementation of competition policy has a positive effect on the retail industry. We have used panel data for 24 countries in transition for a period of 5 years and modeled them with appropriate econometric techniques. The results of our research lead to a conclusion that competition policy has positive effect on the retail industry in countries in transition. We determined that countries which administer competition policy more effectively and efficiently have a more developed retail industry and higher turnover in retail per capita. In spite of that, countries in transition should, in parallel with other areas, develop their competition policy if they wish to adequately implement their retail development policy.

Keywords: *retail, competition policy, transition countries, panel model*

1 Introduction

Trade presents one of the most turbulent economic industries. There are different types of trade. Above all, trade is divided by product type and category meant for trading, food, furniture, machines, tools, appliances, etc. Next, it is divided by type of trade, which includes wholesale and retail. Our further research will be concentrated to retail. Competition policy has its main goal to preserve consumer welfare. It is achieved by preserving a high level of market competition. High level of competition policy includes three areas: control of concentration, misuse of dominant position and restrictive agreements.

The subject of this research is examination of competition policy effects on the retail industry development. In literature there are numerous researches which examine the effects of competition policy on the GDP (Kronthaler and Stephan, 2007), but researches that econometrically model competition policy effects on the retail industry can not be found. Therefore, our base hypothesis in this research is the following:

Efficient implementation of competition policy has positive effect on retail industry development

In order to test this base hypothesis, we will use an econometric model based on panel data. Lately, there has been more and more use of panel data in econometric modeling. Panel data is used in different areas of economy.

The paper is consisted from several thematic parts. At first, a review of relevant literature was created. After that, the information and research methodology used are described. In the third part, an empirical research and testing of base hypothesis was conducted. At the very end of the paper conclusive opinions were prepared.

2 Literature Review

The introduction of competition policy is linked to the Sherman Act of 1980 in USA (Perloff, 1999). The number of countries in which competition policy is implemented has significantly grown for the past years, so at the beginning of the XXI century competition policy is being implemented in over 100 countries in the world (Kronthaler and Stephan, 2007). In world-renown journals related to competition policy there are a great number of researches that confirm the positive effect of competition policy on economic development (Rubinfeld, 2009). Analysis of efficiency of competition policy implementation has a long history in econometric studies (Nicoletti and Scarpetta, 2003). There are a great number of researches which analyze the effect of competition on company and industry performance in general (Aghion et al., 2005).

There are two justifications for competition policy enforcement that have been mostly cited. Firstly, competition policy may be used to correct deviations in competitive behavior. These corrections increase consumer welfare at the expense of manufacturers with potential gains in welfare to society. Secondly, interest groups may lobby for antitrust enforcement to redistribute wealth from producers to their competitors, and other perhaps less efficient producers. In that case, the net impact on society is more likely to be negative (Ghosal and Gallo, 2001). Some authors have examined the related issue of patterns of implementation of competition policy in transition economies (Fingleton et al., 1996; Dutz and Vagliasindi, 2000). We may conclude that there is great literature in which competition policy effect is examined on other variables. However, there still aren't any researches related to the problem and measurement of effects of competition policy on the retail industry. That is why this is the subject of our research.

Researches based on following units of observation in different time intervals are modeled by econometric models based on panel data (Dragutinović-Mitrović, 2002).

In economic literature there are a great number of researches that are actually based on panel models. The pioneers of modern econometric panel data analysis were Mundlak Y. (1978), Zellner A. (1962), Balestra P. and Nerlove M. (1966) who had set the foundations by their work in this field thirty years ago. Great contribution to this area of econometrics was given by Maddala G.S. (1993), Hsiao C. (1986), Baltagi B.(1995), Hausman J.A. (1986) and many others. The econometric interest on panel data has been the result of at least two different types of motivation (Arellano, 2003). At first, the desire to exploit panel data for controlling unobserved time-invariant heterogeneity in cross-sectional models. At second, the use of panel data as a way of disentangling components of variance and estimating transition probabilities among states.

In literature, there are numerous methods of measuring the level of competition policy implementation. They can be divided into three groups (Nicholson, 2008): methods based on surveys, methods based on binary variables and methods based on input and output measurement.

Methods based on surveys enable a comparative analysis of competition policy regimes in different countries, using the experiences of individuals who conduct business activities in those countries. These methods are used for bigger and more complex analyses which are used in a small number of countries. The most significant analysis of this type is conducted within the European Bank for Reconstruction and Development (EBRD). By this analysis, three countries in transition are examined. The result would be an index that measures the level of competition policy implementation. This index is measured upon the following three categories (Duts and Vagliasindi, 2000): enforcement of law related to competition policy, public promotion and advocacy of competition policy and institutional effectiveness of competition policy. EBRD announces each year in its publication *Transition Report* the index of competition policy implementation.

*Global Competition Review*¹ announces its annual survey in which it analyzes efficiency of anti-monopoly institutions in various jurisdictions. Individuals subjected to the survey are usually direct or indirect users of services provided by the regulatory body for competition policy, or those stakeholders whose interests depend on effective and efficient work of regulatory bodies. Above all, they include different business associations, lawyers, consumer associations, economists and other. Questions in the surveys link different terms such as mergers, cartels, technical expertise, autonomy, etc. (Nicholson, 2007).

The World Economic Forum $(WEF)^2$ conducts a survey each year over a great number of countries. Individuals subjected to the survey are business leaders who, through a specific question, value the effectiveness of competition policy in their countries. The value stretches from 1 to 7. Value 1 means that the competition policy is weak and ineffective, while value 7 presents effective competition policy. Results are announced in the publication *Global Competitiveness Report*.

¹ http://www.globalcompetitionreview.com/

² http://www.weforum.org/

In the publication *World Competitiveness Yearbook* announced once a year by the International Institute for Management Development $(IMD)^3$, examinees, mostly people in business, are asked if anti-monopoly law prevents unfair competition in their country. Examinees give answers in values in the range from 1 to 10. Higher value presents better and more effective implementation of anti-monopoly law (Voigt, 2009).

Methods based on binary variables are used in researches which, in structural series, include a great number of countries, while binary variables are used in order to mark the presence or absence of competition policy and law. In econometric models, binary variables also include artificial (*dummy*) variables. In literature, a great number of researches in which this method was used can be found (Evenett, 2003; Kee and Hoekman, 2007).

More complex analyses are based on input measurement, such as the budget of the regulatory body and level of personnel training. These inputs serve as indicators of the level of engaged resources in order to implement competition policy. Effectiveness of competition policy regime can be observed also through output and results, such as number of filed and solved cases before the regulatory body of a certain country (Fingleton et al. 1998). By use of this model, it is possible to conduct high quality and reliable empirical researches. However, a drawback of this model is the fact that it is very difficult to collect such data in a great number of countries.

3 Data and Methodology

3.1 Data upon the research was conducted

The goal of this research is examination of competition policy effects on development of the retail industry in countries in transition. Because of that, we will briefly describe the variables upon which the modeling and examination of base hypothesis was conducted.

For this research, for the purpose of measuring competition policy levels in certain countries, the method based on surveys will be used, so we will consider the remarks made by the European Bank for Reconstruction and Development and in its Transitional report. In these statements, indicators of competition policy implementation level are valued between 1 and 4.33. The lowest level of 1 means that the specific country does not possess a regulatory frame nor a regulatory body for implementing competition policy, while the highest value 4.33 was given to those countries which implement competition policy most efficiently. These indicators reflect the state in a certain country during a one-year period. However, because the process of constructing a modern market economy is a long term process, the effect of efficient implementation of competition policy must be observed for several years. We can not expect the same state on the market in countries whose indicators have reached a certain value, for example 3 in the last year of measurement, and in

³ http://www.imd.org/

countries that have been graded with the same indicator for, let's say, 5 years. It is certain that countries that have had a higher indicator in the past few years have greater influence of competition policy towards market economy development. That is why cumulative observation of indicators is necessary.

Retail presents one of the main business activities of each country and is considered as the trigger of economic development. Great significance of retail is confirmed by the fact that there are a great number of stakeholders. The most important stakeholders in this case are consumers and owners of retail stores, but there are also other stakeholders with strategic expectations and operational demands related to retail development. That is why the level of retail development presents a combination of economic, social, political and physical processes and phenomenon that constantly adapt to each other and to their wider surrounding (Doak 2009).

By having these facts in concern, it is clear that there are several different indicators which directly or indirectly reflect the level of retail development in a certain country. Some of them are (Lovreta 2009; Sun et al. 2008): share of trade in the gross domestic product, share of retail companies in the total number of companies, share of retail industry employees in the total number of employees, turnover in the retail industry per capita, number of retail stores per capita, or number of people per retail store, total area of retail stores per capita, structure of retail store format and other.

If we consider all of the mentioned indicators that can be taken as indicators of retail development, and relative absence of data for a great number of countries, in the research follow-up we will use retail turnover per capita as an indicator of retail development. At official internet websites we have managed to find data about turnover in retail for 24 countries in transition through a period of 5 years, from 2006 to 2010.

3.2 Research Methodology

Because our research will be based on using the panel model, it is necessary to point out the main differences between different types of panel models. The most widely used model in econometric analysis of panel data is the linear model which combines comparable data and time series. In this model, variations of dependent variable *y* are explained by variations of K independent variable, which is the systematic part, and random variations that are not explicitly included into the model, which is the stochastic part of the model.

Now, we will present the general shape for the regression model of panel data:

$$y_{it} = \beta_{1it} x_{1it} + \beta_{2it} x_{2it} + \dots + \beta_{Kit} x_{Kit} + u_{it} = \beta_{1it} + \sum_{k=2}^{K} \beta_{kit} x_{kit} + u_{it}$$
(1)

i=1,...,N; *t*=1,...,T; *k*=1,...,K;

in which symbols are described as below:

 y_{it} - Value of dependent variable for the *i* unit of observation in the *t* period; x_{kit} - Value of *k* independent variable for the *i* unit of observation in the *t* period; β_{kit} - Unknown regressive parameters, which are in the variable by unit of observation and time period in the general form of panel data model

 u_{it} - Random error with the average equal to zero (E(u_{it})=0) and constant common variance (E(u_{it}^2)= σ_u^2 , for each *i* and *t*

For measuring regression parameters of the panel model on the sample of NT data, linear limitations on the parameters are necessary. One of the possible parameter limitations is that all regression parameters are constant, $\beta_{kit} = \beta_k$. Therefore, the model with constant regression parameters has the following form:

$$y_{it} = \beta_1 + \sum_{k=2}^{K} \beta_k x_{kit} + u_{it}, \ i = 1, \dots, N; t = 1, \dots, T.$$
(2)

When all of the regression parameters with independent variables become constant by unit of observation and by time periods ($\beta_{kit} = \beta_k$ for k=2,...,K), and single subparts variable by unit of observations (they include differences between units of observation), then a model of the following form is measured:

$$y_{it} = \beta_{1i} + \sum_{k=2}^{K} \beta_k x_{kit} + u_{it}, i=1,...,N; t=1,...,T;$$
(3)

where $\beta_{1i} = \beta_1 + \mu_i$ is the single subpart of the *i* unit of observation, β_1 is the average value of the single subpart, in which component μ_i shows value of deviation β_{1i} from average β_1 and presents the effects from the model of skipped variables which consider constant values through time, but are different by unit of observation (individual effects). These effects are included in variable single subparts and, in that way, differences (heterogeneity) between units of observation are included in the model.

When it comes to the model with fixed individual effects, individual effects that are included to the single subparts are unknown and fixed parameters. This model is also named as the panel model with artificial variables (LSDV model - *Least squares dummy variable mode*). The model of fixed individual effects can be written in the following form:

$$y_i = \beta_{11}j_1 + \beta_{12}j_2 + \dots + \beta_{1N}j_N + \sum_{k=2}^K \beta_k x_{kit} + u_{it}, \qquad (4)$$

in which $j_1 = \begin{cases} 1 & i=1 \\ 0 & \\ 0$

Variations of dependent variable depend on (K-1) independent variables that consider different values, both by units of observation and time periods, and on variables that are specific for the i unit of observation and constant throughout the observed period.

The model with stochastic individual effects, or the model with random error components (*one-way error components mode*) is used when the single subpart β_{1i} is

treated as a random variable. The hypothesis of this model is that the component μ_i of single subpart ($\beta_{1i} = \beta_1 + \mu_i$) is the stochastic variable, while the other component β_1 is the unknown parameter that presents the average value of single subparts in the whole cardinal set(Matyas and Sevestre, 1996) The model with stochastic individual effects has the following form:

$$y_{it} = \beta_1 + \mu_i + \sum_{k=2}^{K} \beta_k + x_{kit} + u_{it} = \beta_1 + \sum_{k=2}^{K} \beta_k + x_{kit} + v_{it}$$
(5)

or for the *i* unit of observation:

$$y_i = X_i \beta + \mu_i j_t + u_i = X_i \beta + v_i \tag{6}$$

in which $\beta = (\beta_1, \beta_2, \beta_3, ..., \beta_k)$ and $\mu_i + u_{it} = v_{it}$. The term "component of random error" comes from the random error's structure v_{it} . This random error includes two components: 1) random variable μ_i and, 2) error residual u_{it} . The choice of the model measuring method depends on the random error characteristics v_{it} .

Panel models with individual and periodical effects occur when all of the regression parameters and independent variables are constant, while single subparts vary by unit of observation and time periods. This model has the following form:

$$y_{it} = \beta_{1it} + \sum_{k=2}^{K} \beta_k x_{kit} + u_{it} = \beta_1 + \mu_i + \lambda_t + \sum_{k=2}^{K} \beta_k x_{kit} + u_{it}, \qquad i=1,...,N;$$

t=1,...,T. (7)

Single subpart β_{1it} is consisted of three components ($\beta_{1it} = \beta_1 + \mu_i + \lambda_t$). Component λ_t presents the effects of time and include effects of factors that are common for all units of observation in the given period *t*, but their effect on the variations in the dependent variable is different by time periods (1,...,T). If individual and time period effects (μ_i and λ_t) are fixed parameters, then the model would be named model of fixed effects or the model with artificial variables (LSDV model). If we presume that individual and time effects are random variables, the result would be the model with stochastic individual effects. This model is also called the two-factor model with random error components (*two-way error components model*), in spite of its two-dimensional parameter variation (Dragutinović Mitrović, 2002).

3.3 Testing of hypothesis in the panel model

Series of panel data are, above all, concentrated on structure and heterogeneity analysis (variability) between units of observation, but also on analysis of changes in structure through time. Heterogeneity means that the influence of independent variables on the dependent variable varies from unit to unit or through time. Before making the actual choice of a specific model, it is necessary to test if the regressive parameters are constant or variable by units or by time period. Testing is conducted separately for single subparts and parameters with independent variables. Existence of individual and/or time effects can be discovered by tests based on variance 130

decomposition, or by Langrange multiplier tests. If the tests determine the existence of individual and time interval effects, the question of model panel data selection would be raised. Depending which regressive parameters β_{1i} are treated as fixed parameters or as stochastic variables with mutual average value and constant variance, a specific model can be used in the actual research, including the model of fixed effects or the model of stochastic effects.

As the main criteria for choosing between fixed and stochastic effects in the panel model, economics literature most often mentions the nature of sample data and the aim of statistic research (Hsiao, 1986; Hausman, 1978; Mundlak, 1978; Maddala, 1993). For example, if the data is related to certain geographic regions or industries, the model of fixed effects would be suggested. If the units are randomly selected from a great cardinal set, the model of stochastic effects would be used. By use of this model, measurement of a large number of parameters and loss of independent degree number can be avoided (as in the fixed effects model). One of the arguments for using the model of stochastic effects is that the actual effects μ_i should be observed as a random variable, because they are a part of the error v_{it} that is also a random variable itself (Dragutinović Mitrović, 2002).

If the research goal is statistic conclusions about the cardinal set characteristics (for example, variance or average) based on a random sample from the set, then stochastic specification of the model is used, but when the conclusion is determined by individual characteristics of the selected sample, the model of fixed effects is used.

4 Empirical Analysis

We will continue to examine the effects of competition policy on retail market development by applying the econometric model based on panel data. The base specification of the panel model used in our analysis has the following form:

$$y_{it} = \beta_1 + \mu_i + \sum_{k=2}^{K} \beta_k x_{kit} + u_{it} = \beta_1 + \sum_{k=2}^{K} \beta_k x_{kit} + v_{it}$$
(8)

Or for the *i* unit of observation:

$$y_i = X_i \beta + \mu_i j_i + u_i = X_i \beta + v_i \tag{9}$$

This model presents unchangeable regressive parameters with describable variables, while variations per observation unit are included by random error through their components (individual effects μ_i and residuals of random error u_{it}).

Before econometric analysis, a short review of descriptive statistic indicators of observed parameters comes next. Dependent variable in the econometric model is the actual turnover in retail trade per capita (*Real Retail Trade Turnover per Capita - rrttpc*). Independent variables are levels of competition policy efficiency (*comp*) expressed in cumulative indicators given according to EBRD remarks in Transition Reports and GDP per capita (*gdppc*).

At first, we will present correlation analysis of variables observed.

Correlation (Probability)	rrttpc	comp	gdppc
rrttpc	1 (-)		
comp	0.622 (0.000)	1 (-)	
gdppc	0.905 (0.000)	0.575 (0.000)	1 (-)

Table 1. Correlation matrix

From the correlation analysis, we can see that the dependent variable (rrttpc) is highly correlated with the independent variable. The coefficient of correlation between dependent and independent variables is high and statistically significant because p value is equal to 0 for all coefficients, which we had expected. Despite of this fact, we can not make any conclusions based on correlation analysis. It is necessary to find an appropriate econometric model and examine the effects of independent variables on dependent variable in case of their mutual and parallel activity, or effects. At first, it is necessary to consider summarized statistics related to the dependent variable, in this case rrttpc, in order to determine which comparable data includes greater nonsystematic movement of rrttpc based on average value deviation.

CROSSID	Mean	Std. Dev.	Obs.
Mongolia	95.2	18.67351	5
Tajikistan	117.2	14.3248	5
Georgia	160.4	46.72045	5
Ukraine	460.2	72.75782	5
Moldova	533	133.186	5
Macedonia	677.2	71.29306	5
B&H	691.8	110.9085	5
Kazakhstan	783.2	110.339	5
Armenia	943.2	142.1397	5
Serbia	1276.6	224.7405	5
Romania	1123.4	213.1638	5
Azerbaijan	899	358.8029	5
Montenegro	1393.6	239.8818	5
Belarus	1372.2	235.4181	5
Latvia	1912.4	193.5105	5
Lithuania	1897.8	204.4449	5
Bulgaria	1820.6	268.0416	5
Turkey	1748.6	148.776	5
Hungary	2222.4	142.082	5
Estonia	2631.4	209.8161	5

Table 2. Descriptive statistics of dependent variable per observation unit

Russia	2532.4	218.7357	5
SlovakR	3177.2	225.2681	5
Poland	3548.4	430.7375	5
Slovenia	3840.4	422.0039	5
All	1494.075	1068.407	120

Based on statistics $\overline{y_{it}} = \frac{\sum_{t=1}^{T} y_{ijt}}{T}$ average value is determined, so it is confirmed that it reaches highest values in Slovenia, Poland and Slovak Republic. The main conclusion is that during panel model specification individual effects should be

conclusion is that during panel model specification individual effects should be included as a result of various values that consider variables per observation unit. The next table will present descriptive statistics per time periods.

DATEID	Mean	Std. Dev.	Obs.
2006	1262.917	963.5331	24
2007	1439.042	1061.069	24
2008	1675.875	1209.172	24
2009	1505.208	1058.771	24
2010	1587.333	1079.822	24
All	1494.075	1068.407	120

Table 3. Descriptive statistics of dependent variable per time periods

Based on statistics $\overline{y_t} = \frac{\sum_{ij=1}^{N} y_{ijt}}{N}$ annual averages for all countries are determined.

We can conclude that *rrttpc* had average growth of 2,9% (CAGR) in the observed countries.

The procedure of selecting an optimal panel model includes several phases. Firstly, testing of regressive parameters variability is conducted. After that, an optimal model specification is determined. In the third phase, valuation is conducted followed by selection of optimal panel model.

Variability of regressive parameters can be determined upon several tests. We will use the F-test (ANOVA) upon which we will check whether variations per unit observation are statistically significant. It is necessary to add that the mentioned test determines variability of dependent variable, in this case *rrttpc*. We set a base hypothesis that variation of *rrttpc* per observation unit is not significant H_0 : $\mu_i=0$, or variance $\sigma_{\mu}^2 = 0$ (no individual effects), compared to alternative that there are individual effects H_i : $\mu_i \neq 0$, or variance $\sigma_{\mu}^2 \neq 0$. We will begin our analysis by determining existence of individual effects.

Table 4. Results of testing the dependent variable variation per observation unit

Method	df	Value	Probability
Anova F-test	(23, 96)	123.1983	0.0000

The Anova F-test suggests that we should accept the alternative hypothesis related to existence of individual effects and that base zero hypotheses are not valid H_0 : $\mu_i=0$, $\sigma_{\mu}^2=0$. The presented test can be used for testing the existence of time effects, more precisely, heterogeneity of observed variable per time basis. We test the significance of dependent variable variations through time and set a base zero hypothesis $H_0: \mu_t = 0$ which means that there are no time effects, opposite to alternative $H_1: \mu_t \neq 0$ which show existence of time effects.

MethoddfValueProbabilityAnova F-test(4, 115)0.5083340.7297

Table 5. Results of testing the dependent variable variation per observation unit

The Anova F-test suggests that we should accept the base hypothesis of time effect nonexistence. Based on preliminary Anova F-test results, we have determined optimal model specification, which presents model specification with individual effects. Heterogeneity caused by changes in the observation units is the expected outcome of the analysis, in spite of the fact that the mentioned model is more concentrated to the structural dimension.

Specification of individual effects in the model explicitly shows us that the model should not be specified with all constant regressive parameters, including the single subpart, but a more complex question is raised: should the model with individual effects be specified as fixed (fixed effects FE) or together with components of random error (random effects RE). The decision to choose an optimal model will be a consequence of valuating both specifications, RE and FE, including realization of appropriate tests for selection. Firstly, we will estimate the model with fixed effects.

Variable	Coefficients	t- value	Prob.
constant	246,931*	2,206	0.0298
comp	8,967*	2,766	0.0068
gdppc	0,120*	12,320	0.0000
\mathbb{R}^2	0.991		0.0000

Table 5. Results of model estimation with fixed effects

Legend: * significant at significance level of 5%

From the estimated model, we may see that regressive coefficients are positive. From the specification above, it is possible to confirm the existence of individual effects based on the F-test and the ratio credibility test.

Individual Effects Test	Statistic	d.f.	Prob.
Cross-section F	70.204868	(23,94)	0.0000
Cross-section Chi-square	348.024041	23	0.0000

Table 6. Tests of individual effects existence

Both tests confirm the existence of individual effects in the model so the base zero hypothesis that is related to inexistence of individual effects in the model is rejected. It is clear that statistically significant variability *rrttpc* exists in different countries. The next step in the analysis is evaluation of the model with stochastic effects based on the REGLS⁴ method.

Variable	Coefficients	t- value	Prob.
constant	168,516	1.202	0.2319
comp	9,223*	2,959	0.0037
gdppc	0,129*	14,804	0.0000
\mathbb{R}^2	0.733	0.733	

Table 7. Estimation of specification with stochastic individual effects

Legend: * significant at the significance level of 5%

Table 8. Results of component random error variance evaluation

Effects Specification	S.D.	Rho
Cross-section random	450.5844	0.9370
Idiosyncratic random	116.7899	0.0630

Within the previous table, results of random error estimation are shown: evaluated individual effects $\hat{\sigma}_{\mu}$ =450,584 and random error residual $\hat{\sigma}_{u}$ =116,789. Also, through the Rho parameter, a relative share of evaluated elements in the total random error variance is presented and it values 0,937 for individual effects, while for the random error residual the value is 0,063. The presented indicators enable us to calculate ponder Θ .

$$\hat{\theta} = \frac{\hat{\sigma}_{u}^{2}}{T \hat{\sigma}_{\mu} + \hat{\sigma}_{u}} = \frac{116,789^{2}}{5 \cdot 450,584^{2} + 116,789^{2}} = 0,013$$
(10)

The essence of ponder Θ is the capability to transform the model with random error components, which can provide the possibility for evaluating the model based on the REGLS method. As the ponder value does not gravitate to zero, so as the evaluation of individual effects $\hat{\sigma}_{\mu}$ does not equal zero, we can confidently claim that the optimal model is not the model with constant regressive parameters and constant single subpart and that the random error does not include all variations per unit of observation.

Upon evaluation of the model with fixed and stochastic effects, we come across the last phase that includes the actual selection of the optimal panel model. From the stochastic specification, the Hausman test is conducted and it relates to existence of statistically significant difference in the evaluation of model with fixed and stochastic

⁴ Evaluation is conducted by the Swamy-Arora method

individual effects. The presented test also shows if the model possesses endogenous elements, due to disruption of model hypotheses with stochastic effects so that there is no correlation between describable variables and individual effects, that $E(\mu_i X_{ii})=0$ is not valid, or that the hypothesis $E(u_{ii}X_{ii})=0$ is not valid so that the error residual is not correlated with regressors. This test to a certain point includes the problem of heteroscedasticity. We will test the following hypotheses:

 $H_0: E(\mu_i X_{it}) = 0$ Evaluation of the RE model is efficient compared to the covariation grading of the FE model; the RE model is selected.

 $H_1: E(\mu_i X_{it}) \neq 0$ Evaluation of the RE model is biased and inconsistent, the FE model is selected.

Hausman- test	Chi-Sq. Statistic	Chi-Sq. d.f	Prob.
Cross-section random	5,832	2	0,0541

Table 9. Hausman- test

The test result rejects the base zero hypothesis at a level of statistical significance of 10%, which means that there is a statistically significant difference between FE and RE model evaluation. In other words, in the RE model some describable variables are correlated with individual effects (component of random error). In such situation, evaluations of RE models are biased, so the standard procedure would be to choose the FE model because it offers consistent evaluations (co-variation method eliminates individual effects, resulting in disappearance of the problem of their correlation with some of the regressors). Therefore, in this case we choose the model with fixed effects.

The coefficient that stands next to the *comp* variable (Table 5) is positive and statistically significant at the significance level of 5% based on which we may claim that higher level of competition policy has positive effects on the retail industry per capita in countries in transition. The base hypothesis of this research that competition policy has positive impact on retail market development is confirmed. Implementation of the panel model has ensured a thorough form in the context of structural and time dimension. Modeling of turnover in retail per capita in the function of competition policy, but also other relevant variable gdppc, has enabled an increase of given regressive coefficient by maximizing the percentage of describable variations. In the model, the presence of individual effects is confirmed. Existence of individual effects points out the variability of the *rrttpc* by various units of observation. These differences are explained as usual in transitional periods in each of the observed countries. Although each of the 24 analyzed countries characterizes transition, the economic environments before and during the transitional period were inflicted by various tendencies and, therefore, different fluctuations of turnover in retail per capita. At the end we concluded that efficient administration of competition policy has a positive effect on retail market development in the shape of turnover in the retail market per capita.

5 Conclusions

Trade, as one of the most important industries of each and every economy, deserves to be subjected to various types of analysis all over the world. The aspect of trade that deserves most attention is retail, which took over the leading role in the supply chain for the past years. In addition, development of the competition policy is currently undertaken in countries in transition, followed by wider implementation. In this research, the effort was made to examine the effects that competition policy has on retail industry development in countries in transition by using panel model. As a dependent variable, turnover in retail per capita was taken and it reflects development of the retail industry. For independent variables, we have selected the indicators of competition policy implementation level and GDP per capita. Results of this research point out the need for implementing competition policy for the purpose of retail strategy development.

The paper includes a theoretical framework of the panel model with a focus to individual and time effects, including model specification with fixed or stochastic effects. Based on collected data, an optimal panel model was selected upon which we determined the existence of positive effects of competition policy on the retail industry development in countries in transition. At the end of the research, it is necessary to add that conclusions were made upon the base panel models. Further researches should be continued in the form of examining the model specification, functional form and group of observed variables, and also by testing the existence of different types of econometrical problems.

References

- Aghion, P., Bloom, N., Blundell, R., Griffith, R., Howitt, P. (2005) Competition and Innovation: An Inverted-U Relationship, *Quarterly Journal of Economics*, Vol. 120, No. 2, p. 701-728.
- Arellando, M. (2003) Panel Data Econometrics, Oxford Unoversity Press, New York.
- Balestra, P., Nerlove, M. (1966) Pooling cross section and time series data in the estimation of a dynamic model: The Demand for Natural Gas, *Econometrica*, Vol. 34, No. 3, p. 585-612.
- Baltagi, B. H. (1995) *Econometric Analysis of Panel Data*, John Wiley & Sons, London.
- Doak, J., (2009) An inspector calls: Looking at retail development through a sustainability lens, *Journal of Retail & Leisure Property*, No. 8, p. 299-309.
- Dutz, M. A., Vagliasindi, M., (2000) Competition policy implementation in transition economies: An empirical assessment, *European Economic Review*, Vol. 44, p. 762-772.
- Dragutinović-Mitrović R. (2002) Analiza panel serija, Zadužbina Andrejević, Beograd.
- EBRD (2006-2012) Transition report, www.ebrd.com/transitionreport.

- Evenett, S. (2003) Do all networks facilitate international commerce? US law firms and the international market for corporate control, *Journal of Japanese and International Economics*, 17(4), p. 520–537.
- Fingleton, J., Fox, E., Neven, D. J., Seabright, P. (1998) Competition policy and the transformation of Central Europe, CEPR, London.
- Ghosal, V., Gallo, J. (2001) The Cyclical Behavior of the Department of Justice's Antitrust Enforcement Activity, *International Journal of Industrial Organization*, Vol. 19, p. 27-54.
- Hausman, J. A. (1978) Specification tests in econometrics, *Econometrica*, vol. 46, No. 6, p. 1251-1271.
- Hausman, J. A., Griliches, Z. (1986) Errors in variables in panel data, *Journal of Econometrics*, Vol. 31, No. 1, p. 93-118.
- Hsiao, C. (1986) Analysis of Panel Data, Cambridge University Press, London.
- Kee, H.L., Hoekman, B. (2007) Competition Law as Market Discipline?, *European Economic Review*, 51(4), p. 831–858.
- Kronthaler, F., Stephan, J. (2007) Factors accounting for the enactment of a competition law – an empirical analysis, *The Antitrust Bulletin*, Vol. 52, No. 2, p. 137-168.
- Lovreta, S. red. (2009) *Strategija razvoja trgovine Republike Srbije*. CID Ekonomskog fakulteta u Beogradu, Beograd.
- Nicoletti, G., Scarpetta, S. (2003) Regulation, Productivity and Growth: OECD evidence, *Economic Policy*, 18, p. 9-72.
- Maddala, G. S. (1993) *The Econometrics of Panel Data*, Volumes I & II, Edward Elgar Publishing, Cheltenham.
- Matyas, L., Sevestre, P. (1996) *The econometrics of panel data, a handbook of the theory with applications*, Kluwer Academic Publishers, The Netherlands.
- Mundlak, Y. (1978) On the pooling of time series and cross section data, *Econometrica*, No. 57, p. 69-85.
- Nicholson, M. W. (2007), Quantifying Antitrust Regimes, *Erasmus Law and Economic Review 3*, no. 1, p. 41-62.
- Nicholson, M. W. (2008) An Antitrust Law Index for Empirical Analysis of International Competition Policy, *Journal of Competition Law and Economics*, 4(4), p. 1009-1029.
- Perloff, J. (1999) Microeconomics, Addison Wesley, Boston.
- Rubinfeld, D. L. (2009) Evaluating Antitrust Enforcement: Economic Foundations, in Hawk, B. E. (eds), Annual Proceedings of the Fordham Competition Law Institute, Juris Publishing, Inc., Fordham, p. 457-469.
- Sun, L.N., Kaz, R., Chew, M. (2008) Development of retail life cycle: the case of Hong Kong's department store industry, Asia Pacific Business Review, p. 107-121.
- Zellner, A. (1962) An efficient method of estimating seemingly unrelated regression and tests for aggregation bias, *Journal of the American Statistical Association*, 57, p. 348-368.

Veigt, S. (2009) The Economic Effects of Competition Policy: Cross-Country Evidence Using Four New Indicators, *Journal of Development Studies*, Vol. 45, No. 8, p. 1225-1248.

Appendix

Table 1. Assessment of individual effects for each observation unit (FE)

	CROSSID	Effect
Mongolia	1	-631.385
Tajikistan	2	-510.735
Georgia	3	-669.852
Ukraine	4	-500.95
Moldova	5	-167.723
Macedonia	6	-336.65
B&H	7	-335.512
Kazakhstan	8	-618.173
Armenia	9	101.4012
Serbia	10	159.8893
Romania	11	-371.787
Azerbaijan	12	-187.266
Montenegro	13	245.5267
Belarus	14	192.8154
Latvia	15	-150.427
Lithuania	16	-134.33
Bulgaria	17	472.5755
Turkey	18	200.8697
Hungary	19	-105.084
Estonia	20	211.6477
Russia	21	814.1524
Slovak R	22	524.0225
Poland	23	1448.693
Slovenia	24	348.2807

Table 2. Test of auto-correlation in the model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
С	-4.545015	7.503724	-0.605701	0.5462
D(RESIDFE)	0.502640	0.056533	8.891119	0.0000

Table 3. Wald Test of auto-correlation of residual

Test Statistic	Value	df	Probability
F-statistic	79.05199	(1, 94)	0.0000
Chi-square	79.05199	1	0.0000

Building economic institutions and reforming the economy: The case of Kosovo

Linda Tahiri Rukiqi¹

¹South-East European Research Centre (SEERC), International Faculty, The University of Sheffield,

ltahiri@seerc.org

Abstract. War-torn and post-conflict countries face severe development challenges because of weak institutional capacity. In order to start economic activity and facilitate growth, these countries must construct/reconstruct economic institutions. This paper uses the Kosovo case to show the way institutions were built from ground up and the need the country faces to strengthen the medium and long term economic performance. In order to estimate the effects of structural reforms on economic performance, the short paper proposes to firstly measure the impact of up to date real sector reforms, namely reforms in energy and telecommunication sectors, and trade liberalization, on Kosovo's main economic indicators.

Keywords: economic institutions, institutional building, Kosovo, economic reforms

1. Introduction

The importance of institutions for the development of nations has been a topic of numerous economists (see North 1990; Hall and Jones, 1999; Acemoglu, Johnson and Robinson, 2001). War-torn or post-conflict fragile states face severe development challenges such as weak institutional capacity, poor governance and political instability. In order to restart economic activity, facilitate growth, and provide public goods and services, these countries must construct new institutions or reconstruct collapsed ones. Furthermore, as mentioned by Brinkerhoff (2005), in post war/conflict countries, there is a sense of urgency to build or reconstruct their institutions quickly.

As Acemoglu, Johnson and Robinson (2001) put it, economic institutions matter for economic growth because they shape the incentives of key economic actors in society. This short paper uses the case of Kosovo to show the way institutions were built from ground up and the need for reforms the country faces on its road to sustainable development. The following section describes the institutional building in Kosovo, starting from the United Nation's administration of 1999, following the formation of the Kosovar Government, and moving on with economic reforms. Section 3 presents the proposed methodology to measure the impact of up to date structural reforms on economic performance, whereas section 4 concludes.

2. Institutional building in Kosovo

Historically, Kosovo was one of the poorest regions of the former Federal Republic of Yugoslavia and, especially during the 90's, it struggled with high levels of poverty, relatively low growth, and extremely high unemployment. The war/conflict caused Kosovo additional destruction of economic infrastructure, severely damaged bridges, roads and other infrastructure. Furthermore, 40 percent of water sources in Kosovo were contaminated and the energy corporation was far from reaching electricity demand. In 2000, the IMF staff amounted the GDP to 3 billion DEM (1.45 billion USD) with a GDP/capita of 759 USD, and half of the population with consumption levels below the poverty line (IMF, 2001). Coming from decades of neglect, diminished capital, and virtually non-existent economic institutions, institutional building in Kosovo was needed from scratch.

In 1999, the United Nation Mission in Kosovo (UNMIK) used its authority under the UN resolution 1244 to set up a civilian administration in Kosovo as the interim government. UNMIK's mandate included civilian administrative functions, support of reconstruction and economic development, and maintaining of law and order. Following this, in 1999, a Central Fiscal Authority was established in Kosovo and put in charge of budget formulation, implementation and tax policy. The Department of Reconstruction was put in charge for preparing the public investment program and coordinating donor assistance. The inherited trade regime was replaced by a simple one with a single 10 percent tariff rate which was administered by the Customs Department (also newly established). A tax administration with a simple tax system was set up. Furthermore, a Banking and Payment Authority was established to provide a system for domestic payments, license, supervise domestic banks.

A couple of years after the war/conflict ended, following the first Kosovar wide election of November 2001, the three year term assembly was established, the President was elected by the assembly with largely symbolic powers, and the Government was formed. The Kosovar Government, however, had basic functions of administration management only; it was not in charge of economic policy and property management.

In early 2000's, Kosovo, with plenty of room to grow, showed signs of economic growth, improvement of private sector and recovery of the provision of public services. Nevertheless, unemployment and the trade deficit were persistent problems. Since donor money was reducing, an economic policy to develop domestic production was required. According to Acemoglu and Robinson (2012), even in cases of non awareness of policies, policies that increase citizens' welfare could be quickly learned and used; however, the limited power of the Kosovar Government could have also played a role in preventing long-term economic planning orientation for development.

Kosovo's public enterprises, including here the energy corporation, the post and telecommunication company (PTK) and Pristina Airport, needed urgently to be reformed; a state needs public infrastructure to boost economic activity. Being in a severe deteriorated condition because of disinvestment and neglect prior to 1999, Kosovo power plants were in a serious state of despair and posed a hazard to both it's workers and the public. Kosovo citizens were prone to very frequent blackouts due to incapability of the energy corporation to meet electricity demands. Furthermore, the energy corporation had a very low income due to low electricity prices. In this regard, prices of electricity, which were kept artificially low during the Milosevic era, were raised because of the insufficient income to cover the cost of producing, distributing electricity, and maintaining equipment. The energy generation and energy transmission were later separated. With the help of donors, the Energy Regulator was created, the Ministry of Energy and Mining was formed, and the independent electricity transmission system operator, KOSTT, was established.

Although PTK seemed to have been the single source of the sizeable accumulation of deposits in the public enterprise sector during the early 2000's, it suffered from serious internal accounting problems and non transparent financial operations. PTK, a very profitable company, was in a monopolistic position for years. The telecommunication sector needed to be opened up to private sector competition as soon as possible, which was unfortunately prolonged for years, until the private telecom operator IPKO entered the market years later.

In 2007, Kosovo became a part of the CEFTA agreement, a regional multilateral free trade agreement, showing commitment for a liberal orientation of trade policy and willingness for regional cooperation and economic integration.

In early 2007, after the UN Special Envoy, Martti Ahtisaari, presented a comprehensive plan for resolving Kosovo's status, which outlined a framework for Kosovo's independence, Kosovo became a sovereign country in February 2008. That followed a full transition of responsibilities from UN institutions to Kosovo authorities (monitored by an international civilian office (ICO)).

In 2008, Kosovo adopted its own constitution, which defined the economic orientation as a country with a free market economy. Kosovo Privatization Agency was established and continued to manage the transformation of the social enterprises to private property. Furthermore, the inter-ministerial committee was established with the mandate to transform public enterprises such as PTK, Kosovo Energy Corporation and Pristina Airport through involvement of private capital. As a result of the transformation process, Pristina Airport is a successful project of a Public Private Partnership, the distribution and supply of energy is fully privatized and there is an ongoing process for privatizing 75% of the shares of PTK.

Although Kosovo has made significant progress with building social and economic institutions and has experienced robust economic growth in the past decade and a half, it's GDP/capita remains the smallest in Europe, and the trade deficit and unemployment continue to be persistent problems. The table below shows selected economic indicators of year 2012.

GDP (mil.	GDP growth	GDP/ capita	Unemployment	Imports	Exports	FDI inflow	Remittances
E)		(€)		(mil. €)	(mil. €)	(mil. €)	(mil. €)
4,944	2.40% ¹	2,701 ¹	35.10% ²	2,360 ¹	287 ¹	232 ³	605.6 ³

 Table 1.:Economic indicators, 2012

Source: ¹Ministry of Finance of the Republic of Kosovo (2013), ²Kosovo Agency of Statistics (2013), ³Central Bank of the Republic of Kosovo (2013).

In order to improve the performance of economy, pragmatic economic policies and further structural reforms will be required. After years of depending on donor money and remittances, Kosovo, which lacks a Strategy for Economic Development, needs to focus on main issues related to strengthening medium and long term economic development. Sustainable development is considered to be the main vision of further development; therefore sustainable development needs to become an important objective for policy makers (Singh et al., 2009).

3. Proposed methodology

Structural reforms have boosted economic growth and development in other countries through policies that increase the role of market forces and competition in the economy, such as domestic financial sector liberalization, trade liberalization, agricultural sector liberalization. deregulation of the electricity and telecommunication sectors. As stated by Banks (2005), structural reforms typically apply conventional prescriptions for improving growth by removing policy-related distortions and impediments to a well-functioning market economy. Furthermore, IMF (2004) defines them as transformations of institutional frameworks and regulations for markets to function properly.

IMF (2008) divided structural reforms in two main groups: financial sector reforms and real sector reforms, such as reforms in energy, telecommunication, agriculture, and trade reforms. These reforms aim to strengthen the role of market forces and competition in the economy, while maintaining appropriate regulatory frameworks to deal with market failures. Financial sector reforms affect the real sector outcomes, and vice-versa, given the significance of macro-financial linkages inherent in economic performance. Moreover, studies by Henry (2007), Krueger et al. (1992), Sachs and Warner (1995), Dollar and Kraay (2004), McKinnon (1973) show that there is a positive association between real sector structural reforms and economic growth.

Despite Kosovo's progress with building economic institutions and reforming the economic, a Strategy for Economic Development and an economic policy to develop domestic productive capacities are required in order for the country to achieve sustainable development. The country needs to develop instruments for export promotion and FDI attraction, and it needs to focus on sectorial priorities, including continued investment in major projects through the modernization and development of public infrastructure such as energy and telecommunication. The future direction of research is to estimate the impact of real sector reforms, particularly reforms in energy, telecommunication and trade liberalization, on Kosovo's economic performance.

In order to estimate the effects of structural reforms on economic performance, the short paper proposes to measure the impact of up to date real sector reforms on Kosovo's main economic indicators, by using secondary data starting from year 2000. The following variables are proposed to be examined:

- Dependent variables are: 1. real GDP growth; 2. employment rate; 3. FDI inflow; 4. Trade balance.

- Independent variables are structural reforms: trade liberalization (tariff rates), energy (such as the degree of unbundling of generation, transmission, distribution; whether a regulator other than government has been established) and telecommunication (such as competition in local services, whether a regulator other than government has been established).

4. Conclusion

War-torn and post-conflict countries need to construct/reconstruct institutions in order to start economic activity and facilitate growth. Although Kosovo has made significant progress with building economic institutions post-1999, it's GDP/capita remains the smallest in Europe, and the trade deficit and unemployment continue to be persistent problems. In order to improve it's economic performance, Kosovo needs to focus on main issues related to strengthening medium and long term economic development. Structural reforms, including reforms in energy and telecommunication sectors, and trade liberalization, have shown to boost economic growth and development in other countries through policies that increase the role of market forces and competition in the economy. In this regard, in order to estimate the effects of structural reforms on economic performance, the short paper proposes to firstly measure the impact of up to date real sector reforms on Kosovo's main economic indicators.

References

- Acemoglu, D., Johnson, S., and Robinson, J.A., (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. American Economic Review, 91 (5), p. 1369-1401.
- Acemoglu, D. and Robinson, J.A., (2012). Why nations fail: the origins of power, prosperity, and poverty, Crown Publishers, New York.
- Banks, G., (2005). Structural Reform Australian-Style: Lessons for Others? *Presentation to the IMF, World Bank and OECD.* Productivity Commission, Canberra.
- Breinkhoff, D., (2005). Rebuilding Governance in Failed States and Post-Conflict Societies: Core Concepts and Core-Cutting Themes. *Public Administration and Development*, 25 (1), p. 3–14.
- Central Bank of the Republic of Kosovo (2013). *Balance of Payments Report, nr. 12*, Prishtina.
- Dollar, D., and Kraay, A., (2004). Trade, Growth, and Poverty. *Economic Journal*, 114 (1), p. 22–49.
- Hall, R., and Jones, C. I., (1999). Why Do Some Countries Produce So Much More Output per Worker than Others? *Quarterly Journal of Economics*, 114 (1), p. 83-116.
- Henry, P., (2007). Capital Account Liberalization: Theory, Evidence, and Speculation. *Journal of Economic Literature*, 45 (1), p. 887–935.

- International Monetary Fund, (2004). Advancing Structural Reforms, World Economic Outlook, Washington.
- International Monetary Fund, (2008). Structural Reforms and Economic Performance in Advanced and Developing Countries, World Economic Outlook, Washington.
- Kosovo Agency of Statistics (2013). Labour Force Survey 2012: Results for the first half of 2012., Prishtina.
- Krueger, A., (1997). Trade Policy and Economic Development: How We Learn. *The American Economic Review*, 87 (1), p. 1–22.
- McKinnon, R., (1973). *Money and Capital in Economic Development*, Brookings Institution, Washngton.
- Ministry of Finance of the Republic of Kosovo (2013). *Medium Term Expenditure* Framework 2014-2016, Prishtina.
- North, D., (1991). Institutions, Institutional Change, and Economic Performance, Cambridge University Press, New York.
- Sachs, J., and Warner, A., (1995). Economic Convergence and Economic Policies. Brookings Papers in Economic Activity, 1 (1), p. 1–95.
- Singh, R., Murty, H., Gupta, S., and Diksbit, A., (2009). An overview of sustainability assessment methodologies. *Ecological Indicators*, 9 (1), p. 189-217.

Comprehending HRM policies and practices in multinationals within the hospitality sector: country of origin and country of domicile effects

Giovanni Oscar Serafini¹

¹University of Sheffield Management School City College Business Administration & Economics Dept. South East European Research Centre, giserafini@seerc.org

Abstract. This study focuses on human resources management practices and policies applied in hotel multi-national enterprises (MNEs), which belong to a thriving and labor-intensive business area within the wider consumer services industry. Hence, the aim is to assess the determinants of HRM practices in different national contexts within a global luxury hotel chain that manages, franchises, owns and develops branded hotels, resorts and residential and vacation ownership properties around the world. By surveying a case study organization that operates globally within a fast growing service sector it is believed this will enable the exploration of the relationship between corporate guidelines and application of HRM policies and practices to different national contexts, thus contributing to the debate on the nature and application of HRM in the hotel industry. Important findings could occur in relation with the degree of employer-employee interdependence as well as the amount of delegation to employees. Also, as identified in the literature, useful insight will be obtained with regards to the all-important role of HR managers in the sustained implementation of HRM policies and procedures as well as the identification of the barriers that prevent the implementation of HRM in general.

Keywords: *home/host country effects; role of culture; role of institutions; luxury hotel industry.*

1 Introduction

This is a study aiming at analysing the application of Human Resource Management (Mohrman and Worley, 2009) policies in multi-national enterprises in the hospitality industry, while centering on the case of a US hotel chain. In particular, this research work is rooted in the literature on comparative capitalism, and it seeks to shed new light on the nature of institutional effects on firms that cross national boundaries. The focus is on the hotel industry, a rather labour intensive sector, but one that has been relatively neglected in the comparative study of MNE application of HRM policies and practices particularly because it legs behind in their implementation as

compared to other industrial sectors (Lee-Ross and Pryce, 2010; Baum, 2007; Higgins-Desbiolles, 2006; Marchante et al., 2006). Besides, the literature on comparative capitalism has tended to concentrate on the mature markets: this study brings to bear an new body of evidence from emerging markets.

Again, much of the literature on international HRM focuses on comparisons of a limited cross section of countries. In contrast, this study concerns a single organization spanning over 46 countries from both developing and mature markets. Since it is recognized that there is much heterogeneity in the hospitality industry, focussing on a premium chain entails limitations. At the same time, this research allows a very much closer and direct comparison of variations in HRM policy according to locale than would otherwise be the case, within the same organization structure. In other words, it is possible to examine differences between settings within a single and mature organizational context.

More broadly, hospitality services are an essential element of a tourism product related to a destination together with transportation, events, attractions, food and drink. These, in turn, belong to a wider and complex tourism system (Kelliher and Riley, 2002) referred also as tourism market (Cooper et al., 2006) that interacts with the customers through transactions and impacts.

Consequently, as part of tourism markets, the hospitality industry is directly affected by the sheer seasonality of demand of the tourism product: this demand fluctuation often triggered by factors outside the suppliers' control affects the management of human resources within an industry which is particularly labour intensive owing to the nature of the service processes (Reisinger, 2001). The resulting hospitalityrelated employment opportunities create jobs and the income generated has a substantial economic consequence through a "multiplier" effect, especially if a destination relies only on tourism as a main economic activity (Pender, 2005). Thus, by virtue of its link with travel and tourism, the hospitality industry has attained a global profile with an increasing importance in many national economies, thus placing local businesses in direct competition with large multinational hotel chains (Endo, 2006; Pine and Qi, 2004; Mathews, 2000).

Indeed, the hospitality industry represents a main indicator of future trends in terms of business cycle as it follows the development of markets worldwide: as an example, in coincidence with the fact that Asia, with special reference to China and India, the Pacific and Eastern Europe are forecast to become top tourism destinations by 2020 (Johnson and Vanetti, 2005; WTO, 2001), the hospitality industry is growing in those regions at a noteworthy rapid pace. According to the World Travel & Tourism Council, the travel and tourism industry is one of the most important economic sectors of the global economy, as in 2008 it employed over 225 million people around the world and generated 9.6% of global GDP (WTTC, 2009). The World Tourism Organization (WTO, 2001) suggests that the tourism sector has grown from 25 million international tourists in 1950 to 166 million in 1970, to 446 million in 1990 and then reaching a high of 703 million in 2002. After a decline in 2008-2009, 2010 featured a recovery in the international tourism, most notably in developing countries (WTTC, 2011).

It is estimated that over the next ten years (WTTC, 2011), the global travel and tourism economy will increase by 4.3% per year, leading its share of the global economy to just above 10%. In terms of jobs creation, the WTTC estimates that 66

million jobs will be added by 2020. While in 2010 the share of total workforce employed by the travel and tourism economy was 8.1%, this figure is expected to rise by 2020 to 9.2% of total employment. Even by considering the negative effects upon tourism development caused by the 9/11 terrorist attacks, flu disease breakouts, natural calamities, at the beginning of the new millennium a WTO (WTO) report had estimated that there would be 1 billion tourists in 2010, a figure that would increase by 50% within just the next 10 years. This forecast proved right, as 2010 actually demonstrated a strong recovery after the drop of 4% recorded in 2009 caused by the global economic crisis. In real terms, international tourism arrivals in 2010 reached 935 million which represented an increase by almost 7% on the previous year (WTO, 2011).

Although recently more has been written about HRM in the hospitality industry (Davidson et al., 2011), nonetheless this has often occurred by utilizing mainstream HR research frameworks and models (Lucas and Deery, 2004). Further, owing to the vast, complex and multifaceted nature of the hospitality industry, Brotherton (Brotherton) highlights the major challenge faced by academic researchers to appropriately determine the confines and scope of their field of research. Indeed, Guerrier and Deery (Guerrier and Deery) note that most recent research on services industry relate to banking, airline and retail sectors and much less specifically hotels and restaurants.

Consequently, the potential contribution of this PhD lies in its effort to fill in a research void at two distinct levels. First, it is explicitly dedicated to the five-star global hotel industry, thus answers to the two-pronged need of more research in hotel organizations while ensuring clarity of scope within neat research boundaries. Second, following the observation by Drumm (Drumm) that HRM can be described as "some islands of theory in a large sea of untested hypotheses" (p. 46), this PhD study is expected to push further the frontiers of knowledge. In particular, this focussed single-company research feature assists in answering specific research questions (Redman and Mathews, 1998) thus offering a much-needed finer understanding of developments in international HRM (Geringer et al., 2002) applied in the hospitality industry.

2 Literature review

This research work aims to explore and understand commonality and differences in HRM practice within a luxury hotel firm straddling national boundaries. Thus, following an initial overview of the main challenges related to the implementation of HRM by MNEs, the literature review illustrates HRM occurring in the hospitality industry, before moving on to analyse the contemporary trends of thought related to elements affecting HRM in MNEs more broadly. In particular, key current debates reviewed encompass approaches to managing MNE subsidiaries (Holtbrügge and Mohr, 2011), cultural assumptions (Sartorius et al., 2011), institutional theoretical accounts (Brookes et al., 2011) as well as convergence and divergence explanations in relation to management systems within a globalized world (Paik et al., 2011).

2.1 Home and host country effects

MNEs are faced with a key dual challenge involving opposite forces that affect HRM policies and practices implementation: on the one hand, centralization and global integration pushing for standardization, while, on the other, decentralization and local responsiveness pulling towards localization (Rosenzweig, 2006). From the previous sections, it derives that adaptation to local context is an imperative if a transnational firm is to implement a coherent strategy throughout its subsidiaries.

The critical issue is, however, to what extent must a MNE adapt its country-of-origin practices to meet its host-country needs while it is demonstrated that home-country influences exert strong pressures on management systems, the culture of the organization as well as the application of human resource management processes (Lau and Ngo, 1996; Rosenzweig and Nohria, 1994). In particular, Taylor et al. (Taylor and Edgar) indicate that MNEs, by drawing on their home-country competitive advantage, tend to implement homogeneous HRM systems in their subsidiaries overseas. The researchers show that this attempt is especially successful when the gap between the parent and host countries in terms of cultural and legal distance is limited.

According to Cyr and Schneider (Cyr and Schneider), international joint ventures (IJVs) in transition economies of the former Eastern Bloc represent an interesting case in point whereby successful performance is the result of effective co-operation between locals and foreign nationals. As key prerequisites, the researchers indicate that "trust and mutual respect" (p. 219) are instrumental for the attainment of an enduring employee commitment, which, in turn enables the organization to learn and grow while relying on locally sound communication, training, staffing, appraisal and reward systems. Furthermore, in their recent research of Eastern European transition economies, Jindra et al. (Jindra et al.) highlight that beneficial technological diffusion and spillovers towards the development of host-country economies are a direct consequence of a MNE subsidiary role and its technological competence. Again in this specific geographic area, which represents an ideal laboratory to test HRM implementation due to its recent Communist past, Lewis (Lewis) indicates that one of the most precious contribution of MNEs to their subsidiaries and society at large are the training and developmental opportunities offered to the local workforce while instilling the meanings of productivity and quality. As Kostova (Kostova) argues, the transfer of organizational practices within a MNE is effective provided that the social, organizational, and relational dimensions are fulfilled.

Equally important, Ferner and Edwards (Ferner and Edwards) indicate that MNEs should also be considered as structures of power whereby operational outcomes are the effect of a dynamic interplay among resource possession and exchange, formal authority and vested interests. Within this viewpoint, the transfer of best practices is greatly facilitated by coercive comparisons which allow for the exertion of pressures on local management and workforce to align accordingly (Sisson et al., 2003). Nonetheless, Martin and Beaumont (Martin and Beaumont) contend that MNEs should exert a viable pressure on their subsidiaries following the example of a model company such as ABB. More in detail, they remark that internal benchmarking should, on the one hand, consider the local context in terms of culture and institutions

and, on the other hand, the capability and motives of local management to apply best practices.

From a different angle and in contrast with Minbaeva et al. (Minbaeva et al.), Szulanski (Szulanski) indicates that the diffusion of best practice within an organization has to be fostered by eliminating "internal stickiness" (p. 29) caused by elements beyond purely motivational aspects. Indeed, his research leads to the conclusion that barriers to knowledge transfer are caused, firstly, by the inability of the workforce to identify, appreciate and implement new knowledge; then, lack of depth of knowledge; and, last, by ineffective communication and poor interpersonal relationships. As Edwards (Edwards) contends, if there is prevalence of environmental elements favouring synergies and appropriate organizational "maturity" conditions in a global MNE, there may be cases of "reverse diffusion" (p. 696) according to which practices are even transferred from the international subsidiaries to other subsidiaries and domestic plants of the organization.

2.2 Institutions and home country effects

Ferner and Varul (Fournier) argue that HRM is critically affected by the country of origin model of personnel management, which is conditioned by factors within the home national business system. Their research outcome is the result of an analysis of the highly structured German institutional framework determining a considerably reactive and administration-centred personnel management approach pressing against the strategic adaptation of a wider HRM implementation along the lines of the 'Anglo-Saxon' model.

Parent country (ie, country-of-origin) effects represent a category of elements exerting pressures on the way multinationals operate, their organizational culture and, ultimately, their approach to HRM. Thus, MNEs from different home countries apply HRM in different ways for the very reason that, according to Ferner (Ferner):

"even the most global of companies remain deeply rooted in the national business systems of their country of origin (...). Even where the home base does not account for the bulk of sales, operations and employment, the home nation is almost always the primary locus of ownership and control" (p. 19).

As illustrated by Chang et al. (Chang et al.) the pressures exerted onto a MNE in order to be aligned with home country institutional environments represent "push forces" (Brand) and are the result of the MNE degree of embeddedness in its business system (Ferner, 1997). The home country effect can take many forms as demonstrated in numerous studies comparing practices in MNEs from different countries. In particular, nationality of ownership is a critical determinant of MNE business code of conduct and management practices following the comparison of US MNEs as opposed to Japanese and EU MNEs, for instance. While the former tend to be centralized and standardized in the management of HR, Japanese MNEs are more keen to adapt to the local settings (Ferner, 1997). Also, the fact that MNEs mostly gravitate around their home country (Hejazi, 2007) is demonstrated in their concentration of assets, sales generated as well as the marked tendency to fill key positions at the executive level with parent country nationals (Edwards et al., 2007).

The latter, in fact, plays a key role in the transfer process of knowledge and management practices as well as the approach to industrial relations by headquarters towards subsidiaries (Edwards et al., 2007).

Consequently, transfers for organizational development become a prerogative of headquarters nationals (Ondrack, 1985), a persistent ethnocentric focus by MNEs that Mayrhofer and Brewster (Mayrhofer and Brewster) vividly reinforce through comparative survey evidence. Indeed, there are some organizations which are very successful by adopting only parent country policies as happens, for instance, with US oil and gas firms as well Chinese construction MNEs whose operations are almost entirely staffed by expatriates (Corkin et al., 2008; McSherry, 2006; Wood, 2004). While in the former home-country expatriates are vital to ensure the skilful execution of highly technologically sophisticated operations, the latter are extremely successful as compared to other competitors owing to their low-cost but solidly trained and committed Chinese expatriate labor force. This organizational reality illustrates that ethnocentricity is one of the available avenues towards attaining convergence and internal consistency on a global scale (Lau and Ngo, 2001). Indeed, research highlights that MNEs are affected by parent country pressures towards a dominant global model on several key fronts, namely management system, corporate culture and HRM practices (Schuler and Tarique, 2007; Schuler and Rogovsky, 1998; Rosenzweig and Nohria, 1994).

In summary, it is critical to recognize the link parent country effects have with the issue of institutions as previously reviewed. In fact, as firms are located in the centre of a composite interaction web, their effective relationship with "producer groups, employees and other firms" (Hall and Thelen, 2009) is directly affected by the support of institutions from a political economy standpoint. As a result and in proportion to its strength, also the parent country institutional framework influences the approach businesses have to their international operations.

2.3 Organizational isomorphism

The implementation of HRM by a MNE is not a straightforward process, but rather reflects the modus vivendi of subsidiaries and their dynamic interplay vis-à-vis the organization headquarters. In fact, Rosenzweig and Nohria (Rosenzweig and Nohria) in studying human resource management practices in 249 U.S. affiliates of foreignbased MNEs from Canada, Japan and Europe demonstrate that, broadly, affiliate HRM practices are strongly aligned with local practices while, however, differing in particular practices. The extent of "local isomorphism" (p. 241) as the authors put it, is determined by the founding method, the reliance on local inputs, the ratio of expatriates within the workforce of the subsidiary, the closeness of communication with headquarters and the need for internal organizational consistency. According to DiMaggio and Powell (DiMaggio and Powell, 1983) there are three main forms of isomorphic outcomes, each determined by particular processes: firstly, "coercive isomorphism" relates to the pressures onto an organization as occurs through the imposition of law, for instance; secondly, "mimetic isomorphism" indicates efforts of an organization to copy successful practices adopted by other organizations; and, last, "normative isomorphism" describes the adjustment of an organization to what is

regarded as appropriate in a particular environment (Kostova and Roth, 2002; Haveman, 1993).

Notwithstanding, Jackson and Schuler (Jackson and Schuler) highlight that local environment and culture have a key role in the application of HRM systems, policies and practices. On this same line, a study by Beechler and Yang (Beechler and Yang) on Japanese subsidiaries operating in the USA shows that local circumstances decisively influence a MNE's ability to transfer parent-country practices abroad. As a result of particular local constraints, businesses are mostly unable to accurately replicate home countries' practices in their international subsidiaries and thus opt for a locally isomorphic approach (Rosenzweig and Singh, 1991). Indeed, even between similar national business systems such as the American and the British, the transfer of HR practices from the US parent company to the UK subsidiary is impeded by local stakeholders who prevent a MNE to fully implement a corporate diversity policy, as shown in a case analysis by Ferner et al. (Ferner et al.).

2.4 HRM in the hospitality industry

The hospitality industry, which is intertwined with the wider tourism industry, is one of the largest and most dynamic industrial sectors worldwide. In particular, the hospitality industry is identified with all businesses providing, according to Brotherton (Brotherton and Adler):

"a contemporaneous human exchange, which is voluntarily entered into, and designed to enhance the mutual wellbeing of the parties concerned through the provision of accommodation and food or drink." (p. 168)

The expansion of tourism and hospitality industry globally is accompanied by two specific trends: on the one hand, consumers expect more quality and variety in the services and products they purchase while, on the other hand, competition among firms is heated up both nationally and internationally. Consequently, superiority in this industrial sector is attained by businesses which are able to surpass competition in terms of service quality, customer satisfaction and business performance (Augustyn and Ho, 1998). Thus, a key pre-requisite is the implementation of sound HRM practices owing to the fact that hospitality critically depends on successful human interactions (Davidson et al., 2011).

In order to clarify the industrial sector under consideration, there is the need to identify the different business typologies. In fact, hospitality organizations share a most common services framework whereby accommodation is compounded with a number of supporting facilities such as specialty restaurants, shopping arcade, wellness center, which can escalate to a very sophisticated range of services depending on the competition and the market addressed (King, 1995). For the purposes of this PhD, emphasis is given to MNEs operating in the luxury segment of the industry, otherwise referred to as "international hotel chains" which offer product-branded services on a standardized global basis (Whitla et al., 2007).

2.5 HRM in the luxury hospitality industry MNEs

Same as occurs in any other industrial sector, hospitality firms do have the strategic option to expand their business operations internationally, or even globally. Contractor and Kundu (Contractor and Kundu) found that the entry mode into new markets is determined by country-, environment- and firm-specific variables. According to the researchers, 65.5% of hotel international operations are contracted through non-equity agreements, such as franchise agreements, management contracts and strategic alliances. Consequently, these arrangements enable hotel organizations with successful and unique service brands to expand globally. Further, a most recent study by Graf (2009) clearly indicates that stock markets react abnormally whenever the entry mode of the hotel chain fits with the specificities of the host country. In particular, the author found that investors react very favourably to new management contracts in developing countries and new franchise agreements in developed countries, also because of the anticipated benefits resulting from the hotel firms' strategic orientation, contractual control and formality of business operations (Yan et Notwithstanding, the successful implementation of HRM policies and al., 2007). practices in overseas subsidiaries depends on the sound interaction between the owners who do the investment and the hotel management company which either runs the business on their behalf through a management contract or allows the hotel to operate under a franchising agreement (Gannon et al., 2010).

For luxury hotel chains, emphasis on quality throughout is of utmost importance and, therefore, HRM plays a strategic role in strengthening this vision within the organization (Maxwell et al., 2004). In 1999 Hilton, for instance, introduced a worldwide customer service quality initiative named 'Equilibrium' which prompted HRM to develop a strategy branded 'Esprit' that encompassed, beyond a successful staffing function, policies and practices targeting specifically the areas of assessment, compensation, benefits, recognition, career tracking and discipline (Maxwell and Lyle, 2002). Other hotel chains, have put HRM at centre stage with regards to their business strategy along with marketing and operations as in the case of Accor (Aung, 2000), or, even, have achieved remarkable talent management through the dynamic involvement and support of the hotel General Manager to a critical HR area of responsibility (Yeung, 2006).

Nonetheless, despite the global expansion of luxury hotel chains and the subsequent attempt to globalize the HRM function, a closer examination of local contexts demonstrates the existence of unique challenges pertaining to the political, national and cultural settings, which requires the local hotel HR professional to adapt accordingly (Naama et al., 2008; Costa, 2004; Lu and Chiang, 2003). Indeed, Boxall (Boxall) argues that differences in workforce capability, labour productivity and employment systems affect the management of human resources which needs to be modified from country to country.

Furthermore, labor markets critically affect the quality of service professionals: to this end, governments do contribute to the establishment of MNEs through, among others, the education of local nationals in hospitality vocational professions (Baum and Szivas, 2008; Kusluvan and Karamustafa, 2001). Even in developed countries such as Korea, research has shown that Training Managers working in international chain hotels modify training programs in order to make them fit to the local culture

(Lim and Wentling, 1998). Equally important, labour shortage is a ubiquitous challenge to be found even in populous China, however with its distinct peculiarities related to lack of qualified labour force and the unwillingness for university graduates to join the industry (Zhang and Wu, 2004). On the other hand, surveys by Magnini and Honeycutt (Magnini and Honeycutt) and Shay and Baack (Shay and Baack) demonstrate the high failure rate of expatriated hotel managers as well as discrepancies between the expatriated managers' self-perception of performance and subordinate-rated managerial effectiveness.

In conclusion, local specificities compel HRM to adjust to the 'think global, act local' perspective whereby central offices originate wide-scope HRM policies and procedures and then it falls upon the local hotel HR executive to decide and act the best way they see fit to the local context which, incidentally, they know better than anybody else in the HRM function (Enz, 2009). An example of such an approach is offered by Zuehl and Sherwyn (Zuehl and Sherwyn) who contend that, after analyzing employment termination practices in a sample of countries, MNEs feature common employment termination policies presented as general guidelines to be considered, but then empower the local HR executive to decide accordingly.

2.6 Differences in HRM implementation: hospitality MNEs v. local hotels

Although rising, the number of multinational organizations yet represent the sheer minority of all businesses operating in the hospitality industry: in fact, local small and medium sized hotel enterprises (SMEs) prevail all the world over as documented by Baum (Baum) according to whom 90% of hospitality businesses globally belong to the SME type. As a result of their intrinsic lack of professional and business skills, insufficiency of funds and unstable business performance, most SMEs are evidently disadvantaged when it comes to competing on human resources management grounds with larger organizations (Baum, 1999).

Thus, the hospitality industry is theatre to a most striking contradiction: despite the fact that theoretical propositions and empirical evidences clearly indicate that HRM effectiveness is key to service quality, customer satisfaction, business sustainability and profitability as well as competitive advantage, still the reality of many local hotel organizations in terms of workforce management practices and employment conditions often demonstrate the lack of sound HRM (Cho et al., 2006). In fact it is worrying that in practice the role of HRM within such organizations is given low priority compared to other business functions such as finance and sales and marketing (McEvoy, 1984; Rutherford and O'Fallon, 2007; Haynes and Fryer, 2000; Kelliher and Johnson, 1997). This leads to questioning the credibility of the fundamental assumption made and expressed by hotel management according to which employees are the most important asset and key to success in relation to service quality, consumer satisfaction as well as business performance (Losekoot et al., 2001; Maher, 1993).

The reasons behind this inconsistency are varied, and can be traced to the labourintensive character of the industry with a prevalent focus on external labour markets related to easily replaceable low-paid marginal workforce (Davidson et al., 2006; Bohle et al., 2004; Robson et al., 1999). This, coupled with high costs and thin profit margins in a markedly seasonal business environment (Jolliffe and Farnsworth, 2003), causes the sheer divide between theory and reality. It is no surprise, therefore, that the hospitality industry SMEs due to the fact that most often they are not a source of permanent and stable employment, offer uncompetitive retribution for working in anti-social hours while featuring few opportunities for job enrichment and career advancement have a diffused poor reputation, particularly among hospitality and tourism students (Baum, 2002; Kusluvan and Kusluvan, 2000; Choy, 1995; Wood, 1992).

On the other hand, hospitality MNEs with an international or, even, global reach have made remarkable efforts to narrow the gap between words and deeds and upgrade the profile of the industry by drawing from the long-term benefit of sound HRM practices implementation (Bohdanowicz and Zientara, 2009; Rowley and Purcell, 2001). As Baum (Baum) maintains, it is vital for hospitality organizations to develop their workforce so as to render them "more flexible and adaptive to constant change" (p. 133) at both operational and managerial levels. Thus, Ford and Heaton (Ford and Heaton) contend that, due to the considerable intangible aspect of the service experience (Lashley, 1998), the hospitality industry has managed to develop unique competencies in service provision that businesses in any other industrial sectors may consider worth implementing. This because while the hotel industry is particularly capital- and labor-intensive, its logistics and supply chain processes can be as sophisticated as in manufacturing businesses (Dimou et al., 2003) contrarily to other sectors of the services industry, such as marketing and consulting.

Exactly because the hospitality industry is focused on providing service excellence in an extremely competitive environment (Dubéa and Renaghanb, 1999), hospitality MNEs aim to build a solid culture of service so as to guide every employee behavior by having the customer in mind and ensure that their experience is managed successfully (Teare, 1995). In turn, this reflects on managerial skills of MNE hospitality professionals which differ substantially from those required in the manufacturing industry. Particularly, hospitality service delivery systems demand managers to develop special competencies in areas such as organizing, staffing and commanding (Bowen and Ford, 2004).

Consequently, being customer satisfaction and service delivery consistency the main strategic challenges faced by any hospitality concern, staffing is a top HRM priority with the goal to hire employees featuring a strong service attitude (Ladhari, 2009). Nonetheless, taking the local UK hotel market as an example, a survey by Price (Price) highlights that, despite these recruitment efforts, still personnel practices are poor for local hospitality SMEs in the mid-nineties as they were in the mid-eighties as highlighted by Kelliher and Johnson (Kelliher and Johnson), a worrying reality further confirmed by McGunnigle and Jameson (Fournier). However, in the mid-nineties that same local market indicates that large and foreign-owned hotels are featuring a more advanced approach to personnel management (Kelliher and Johnson, 1997), even if, overall, as compared with other business sectors in Great Britain, the hospitality industry features a remarkable divergence in the application of HRM. This is particularly highlighted by the markedly "retaining control/cost control" approach (Lucas, 2002) to management leading to a very hard HRM in practice causing employee dissatisfaction.

Another element worth considering is that the hospitality industry usually features a low level of unionization, whereby sound HRM strategy formulation and

implementation depends on the effectiveness of management (Tanke, 2001). For instance, in the UK Lucas (1996) reports a decline in union membership in the 1980s from 6% to 3%, while union membership among hospitality workers in North America totals 14% (Tanke, 2001). The minor role of unions in the hospitality industry is due (Piso, 1999) to, first, the fact that the overwhelming majority of hospitality businesses belong to the national SME category characterized by a paternalistic style of management (Rowden, 1995); second, the varied and fragmented nature of hospitality industry (Ingram, 1999; Harrington and Akehurst, 1996); then, the very workforce composition does not facilitate the emergence of unions because of dissimilar employee profiles (young students, women, transient workers, minorities), job contracts (part-time contracts, internships, outsourcing and subcontracting practices) (Soltani and Wilkinson, 2010); and last, weak internal job markets (Deery and Jago, 2002) as well as considerably high turnover rates (ILO, 2001; Price, 1994).

In the final analysis, as the research by Wilton (Wilton) reveals, the efficient and effective application of formal and strategic approaches to HRM is directly proportional to the size of the hospitality organization. This observation further strengthens the initial position that hospitality MNEs, owing to their vast resources and scale, are at a net advantage compared to national hotel SMEs in relation with the systematic implementation of sound and innovative HRM practices (McPhail and Fisher, 2008).

2.7 Peculiarities of employment in the hospitality industry

In the hospitality service provision, the relationship between the customer and the employee is asymmetrical (Shamir, 1980) because it is the role of the latter to offer and of the former to receive, provided they want to (Jones and Sasser, 1995). The exchange is based on the fact that the guest must pay for the services with the understanding that they would not harm or disrupt the hotel, other customers as well as the service providers (Radolovic, 2010). There are three key aspects to the hostguest relationship: first, the service provider does not hold a superior position in terms of a particular knowledge vis-à-vis the receiver as would happen with a doctor, an engineer or an economist; second, often the host is economically more affluent and of higher status than the service provider, something which is evident in luxury hospitality establishments (Goeldner and Ritchie, 2006); third, especially in customercontact department, tipping is an important source of income to the employees, something that gives the customer control of their behavior to a certain degree (Azar, These basic elements plus the existence of intense competition put the 2007). customer in a clear position of advantage and control that presses hospitality organizations to forge in all employees an attitude of total customer satisfaction (Heart, 1988).

The main challenge faced by a hospitality professional are the many roles they need to cover at once (Dev and Olsen, 1989): on the one hand, while they aim to serve the needs of the customer they have still to focus on the control of the transaction so as to do it effectively; on the other, the hospitality professional needs to perform according to the organization brand standards, policies and procedures while maintaining their personal emotional balance and self-esteem.

Consequently, the function of human resources in the actualization and delivery of successful hospitality service experiences should not be underestimated: through effective HRM implementation as well as the nurturing of a people-centred organizational culture focussing on the respect of both the external and internal customers (employees), a hotel organization should be truly hospitable. However, the issue still at stake is whether theory is applied in practice (Wilton, 2006).

2.8 Labour intensity

Boddewyn et al. (Boddewyn et al.) recognized the importance of service MNEs as distinct from those non-service which deal with product manufacturing on the grounds that "by the year 2000, more than half of the world's multinational enterprises will be in services" (p. 54). Indeed, about twenty years later, a report by UNCTAD (UNCTAD) proved such forecast right, indicating the ascending role of workforce in the delivery of MNEs. Nonetheless, Lowell (Lowell) vividly contends that:

"World-class talent has enormous scale effects as interaction costs fall because it can be leveraged across ever greater geography. It underlies most opportunities to specialize. It is the ultimate source of all intangible capital. It underpins all the other components of intangible capital necessary for global success. And, unfortunately, it is in short supply" (p. 182).

It is a fact that the hospitality industry critically relies on its workforce as the products and services it offers are pervaded by the human element (King, 1995). As a result, compared to other industries, it is widely recognized that the hospitality industry features high labour intensity which impedes radical downsizing as occurs in other sectors such as manufacturing and agriculture, because of the irreplaceable function of service personnel and the resistance to the application of technology (Nankervis, 2000). Further, the employment relationship between the hotel and its workforce is complicated by the very characteristics of the hospitality service: intangibility, perishability, heterogeneity, variability, simultaneity of production and consumption, and inseparability (Reisinger, 2001). Nonetheless, experts maintain that the level of labour intensity is also affected by the extent of tourism development of a particular destination and its infrastructure as well as the type of business (Baum, 2006).

Consequently, even if at varying degrees, labour expenses by themselves represent a considerable share of the total production costs all over the world in an industry with high cyclicality and tight profit margins (Woodworth and Mandelbaum, 2010). This results in a compression of the remuneration of employees to levels below industry average and in the implementation of practices to contain, if not reduce, payroll cost. Thus, hotel companies often resort to outsourcing services in order to seek a more just-in-time labour provision (Gonzalez et al., 2011; Lamminmaki, 2011) and the application of part-time work contract (Cairncross and Kelly, 2008).

2.9 Employee turnover challenge

The above observations suggest a most challenging reality which is represented by the instability of labour resulting in alarmingly high turnover rates especially at the lower In fact, some researchers have even commented that the ranks (Wood, 1992). hospitality industry features a "labour turnover culture" (Davidson et al., 2010). Indeed, an International Labour Organization report on the hotel, catering and tourism sector (ILO, 2001) shows that in 1997 the U.S. feature a 51.7% turnover rate for line employees compared to 30% in Asia and 42% in the UK. To illustrate how this compares with other industrial sectors, in the case of the UK, hospitality industry turnover is second only to retail trade with 43.5% and far ahead of construction with 25% (ILO, 2001). The ILO additionally highlights that on average 25% of the total workforce is on a part time contract in 15 EU countries, while the figure rises to about 50% for Northern countries such as the UK, Denmark and the Netherlands. As Iverson and Deery (Iverson and Deery) maintain, this overarching turnover culture is further exacerbated by sheer seasonal fluctuations of tourism demand as well as unforeseen and uncontrollable circumstances of social, economic and political nature. Interestingly, as the hospitality industry offers employment ranging from the unskilled to the highly skilled, its workforce composition features a contrast between employees at the periphery against those at the core (Brien, 2010; Knox, 2010). This particular differentiation of workforce depending on their indispensability is reflected on their status within the organization as well as on their job satisfaction which, in turn, is mirrored in their relative turnover rates. In fact, the ILO (ILO) statistics confirm that in 1997, turnover rates in the US lodging industry for supervisors and property managers are about a fifth of those related to line employees as noted before. Further, the research of Lai et al. (Lai et al.) Indicates that peripheral hotel employees are often considered costs rather than resources. Thus, it can be inferred that the difference between a hotel professional and an employee is to be found in the fact that the former, as opposed to the latter, identify themselves in their profession and enjoy more stable working conditions while the latter is exposed to contingent labour arrangements owing to their disposability (Hjalager and Andersen, 2001).

Nonetheless, the hospitality industry (Hui and Hsin-Wei, 2009) has its own share of responsibility in leading turnover to such high levels because of the widespread poor working conditions and human resources practices. This includes unattractive retribution, absence of career planning, uncertainty of employment as well as long and unsocial working hours (Boella and Goss-Turner, 2005) particularly in local hotel businesses that may even reach Orwellian connotations in certain cases (Orwell, 1933). The result is that many qualified employees, after a short stint in an industry featuring mostly seasonal and part-time employment, seek better professional opportunities elsewhere in the labour market. Thus, hotels are especially attractive employers to the "secondary labour market" (Price, 1994), who are not distinctly dedicated to a specific industry such as students, housewives, school leavers and unqualified workers. These job-seekers search for a convenient source of income on a part-time basis (Walsh, 1990) who, while having limited alternative occupational opportunities, appreciate working in a dynamic environment in contact with people (Curtis and Lucas, 2001).

As can be inferred from above, literature documentation (Deery and Shaw, 1999) indicates that employee turnover has both positive and negative consequences. On the positive side, it can be argued that it offers manning flexibility to serve effectively a fluctuating customer demand (Graham and Lennon, 2002). On the negative side, uncontrollably high turnover levels cause stress, damage employee morale, raise the workload to incumbent workforce due to systemic staff shortages and skills, impede the development and improvement of products and services, increase replacement costs, eliminate the benefits of training, with the end result of fostering a disengaged employee attitude leading to loss of revenue and profitability (Lee and Way, 2010).

2.10 Nature of hospitality jobs and related labour markets

Statistics reveal that an overwhelming share of jobs in the hospitality industry is unskilled or semi-skilled and that approximately 36% of jobs relate to managerial, supervisory and skilled (craft) positions (Riley, 1996). This reality has favoured society's widespread perception that jobs in hospitality are of low quality resulting in a downward pressure in terms of compensation and working conditions (Simons and Enz, 1995). Given the practical nature of the business, managers in the hospitality industry give emphasis on the operational skills and on-the-job training capability of the workforce, something that does not necessarily demand particular qualifications (ILO, 2001). Thus, such circumstances featuring both de-skilling and standardization in the name of a McDonaldization of processes (Lashley, 2008; Ritzer, 1998) facilitate the transferability of skills within the industry which justify also the high employee mobility within the industry (Deery and Jago, 2002).

At this point a clear distinction needs to be made between employment needs in developing as opposed to developed countries (Baum, 1996). In developing countries, hospitality industry is compelled to recruit at considerably higher level than it would occur in developed countries owing to the shortage of local workforce who is educated and exposed to the outward culture of tourism and hospitality. The very need to fill the gap related to cultural and communication skills combined with the imperative to ensure world-class service standards, oblige industry operators in developing countries to raise the workforce profile by recruiting overseas. The result is that, according to Baum (1996), in many developing countries the ratio of semi-skilled or unskilled jobs is reduced to 15-25 % of the workforce in the hospitality industry.

By taking into consideration the high labour intensity of the hospitality industry, its considerable turnover rate as well as the critical share of unskilled and semi-skilled jobs over total workforce, it can be argued that the hospitality industry is generally characterized by a weak internal labour market. In fact, the particular nature of the business prevents the majority of establishments to foster a well-organized work environment with clear standards in terms of job description, performance evaluation, pay and promotion. According to research (Woods, 1999), employment in the hospitality industry is especially sought by women thus determining a feminization of the work whereby they would cover mainly mothering roles (housekeepers, waiting staff) or glamorous roles (front desk positions, sales and marketing associates). Nonetheless, developing countries may not feature a pronounced workforce feminization in their hospitality industry due to cultural reasons and women's position

within the society according to which it is custom to have them work behind the scenes or in a back-of-the-house position (Burrell et al., 1997). Consequently, it can be argued that the status of employment in the hospitality industry is affected by the general economic and social settings of a country or a region, such as the development stage, the degree of economic diversity as well as the level of dependence on the tourism activity of an area (Matias et al., 2009).

3 Methodology

The company under consideration is a global deluxe hotel chain that, for the purpose of safeguarding anonymity, will be referred to as HotelCo. This offers a unique research opportunity to understand the influence exerted by the MNE centre towards the periphery in an industry where MNEs are more spatially bound than many other In fact, the peculiarity of international hotel chains is that they cannot, employers. unlike manufacturing and some services such as call centres, move to the locations of lowest possible cost in a race-to-the-bottom strategic choice as previously explored (Contractor et al., 2010; Wright et al., 2005). Rather, they have to base their decisions primarily on consumer demand, resulting in hotel chains being present in some extremely high wage and heavily regulated contexts (Contrepois and Jefferys, 2010; Knox, 2010; McDowell et al., 2007). This makes hotels a very interesting case because they cannot simply exit to avoid regulation as a blanket policy since they need to be where customers are (albeit that they may choose to leave a specific market if the costs of labour, for example, are so great as to make for uncompetitiveness), but rather optimize their strategy according to the particular context and competition settings (Timo and Davidson, 2005).

Thus, the case study organization is a diversified global US hospitality concern with more than fifty years of history, acting as the owner-manager, franchiser and management company of international deluxe hotels, resorts, and mixed-use real estate developments. The HotelCo global mission is to provide authentic hospitality and its goal is to be recognized as the most preferred brand in each market segment it operates while being faithful to a set of core values featured in its culture. In particular, as of March 31, 2013, HotelCo's worldwide portfolio consisted of 508 properties grouped under the EAME (Europe-Africa-Middle East), ASPAC (Asia-Pacific) and LATAM (Latin America) divisions which are coordinated by divisional offices, respectively.

Therefore, the aim of this PhD is to examine the determinants of HR practices within a MNE operating in the hospitality industry and attain new insights into the relative impact of host and home country effects. With the premise that the research aims at surveying a leading global hospitality company operating in 46 countries, the consideration of the national archetypal business systems as enunciated by Whitley (Whitley) is instrumental in order to group national contexts according to their distinctive type of coordination and control systems. In particular, Whitley's model (Whitley) includes: fragmented, coordinated-industrial-district, compartmentalized (i.e. 'Anglo-American model'), state-organized, collaborative ('the Rhineland model') and coordinated ('the Japanese model') business systems.

In the final analysis, using a case study organization that operates globally within the fast growing service sector it is believed it will enable the exploration of the relationship between corporate guidelines and application of HRM policies and practices to different national contexts, thus contributing to the debate on the nature and application of HRM in the hotel industry.

4 Empirical analysis

Since the overall aim of this PhD research is to assess the nature and determinants of HR practices of a US hotel MNE in its foreign subsidiaries from the viewpoint of HR executives, it is necessary to attain the following research objectives:

- i) to identify, within the context of the luxury hospitality industry, the degree of autonomy overseas subsidiary HR executives have in setting their HR policies versus domestic HR executives;
- ii) To compare and contrast the HR practices encountered in overseas subsidiaries versus those in domestic hotels in the case study firm;
- iii) To explore the wider theoretical implications of any homogeneity or diversity within the parameters of business systems theory, and the lessons for applied HR policy and practice.

A quantitative survey method will be used to examine the approaches and implementation of HRM policies and practices in a global luxury hotel chain. As Saunders et al.(Saunders et al.) point out, the quantitative survey "is a popular and common strategy in business and management research", and surveys enable "easy comparisons and appear to be authoritative" (p. 92). Thus, the survey is a useful tool for analyzing and assessing HR practices in a systematic way that allows the researcher to gather similar data and draw useful conclusions following comparisons.

More in detail, respondents are asked to answer questions divided into the following sections:

- A. Autonomy in relation with HRM decisions
- B. Communication and information flows
- C. Employer-employee interdependence
- D. Delegation to employees
- E. Host-country specific factors determining the application of human resource management practices in multinational hotel chain subsidiaries:
 - a. Selection practices
 - b. Compensation practices
 - c. Performance appraisal practices
- F. About expatriates
- G. Workplace information & demographics

Information is sought in order to assess the extent of formal corporate HRM strategy application by subsidiaries, drawing from the input via questionnaire administration to HR executives heading the HR function at overseas subsidiary level

in 7 properties located in EAME (Europe-Africa-Middle East). The fact that participants to the survey are Human Resources executives should provide a unique insight into the process directly from the management representatives in charge of implementing HRM policies and practices at hotel unit level. The wealth of information obtained allows for direct comparison across subsidiaries (Pudelko and Harzing, 2007). Additionally, data gathering is then enriched through semi-structured follow-up interviews.

5 Conclusions

This article aims at laying the theoretical grounds to an ensuing research work in the field of HRM applied to the global luxury hotel industry. The review of literature leads to the conclusion that business and context specificities including host country effects, the role of culture and institutions plus globalization pressures affect the application of HRM in practice across countries where a MNE operates.

Being the hotel group a global organization, the research will offer useful insight on how people management is occurring in different countries and contexts thus allowing the interpretation of results across locations and economies.

References

- Augustyn, M. and Ho, S. K. (1998) Service quality and tourism. Journal of Travel Research, 37 (1), p. 71-75.
- Aung, M. (2000) The Accor multinational hotel chain in an emerging market: through the lens of the core competency concept. *The Service Industries Journal*, 20 (3), p. 43-60.
- Azar, O. H. (2007) Why pay extra? Tipping and the importance of social norms and feelings in economic theory. *Journal of Socio-Economics*, 36 (2), p. 250-265.
- Baum, T. (1996) Unskilled work and the hospitality industry: myth or reality? International Journal of Hospitality Management, 15 (3), p. 207-209.
- Baum, T. (1999) Human resource management in tourism's small business sector: policy dimensions. In: Lee-Ross, D. (ed.) HRM in tourism and hospitality: international perspectives on small to medium sized enterprises. London, Cassell, p. 3-16.
- Baum, T. (2002) Skills training in the hospitality sector: a review of issues. Journal of Vocational Education & Training, 54 (3), p. 343-363.
- Baum, T. (2006) The future of work and employment in tourism. In: Buhalis, D. and Costa, C. (eds.) Tourism management dynamics: trends, management and tools. Oxford, Elsevier Butterworth-Heinemann, p. 130-136.
- Baum, T. (2007) Human resources in tourism: Still waiting for change. *Tourism Management*, 28 (6), p. 1383-1399.
- Baum, T. and Szivas, E. (2008) HRD in tourism: a role for government? *Tourism Management*, 29 (4), p. 783-794.

- Beechler, S. and Yang, J. Z. (1994) The transfer of Japanese-style management to American subsidiaries: contingencies, constraints, and competencies. *Journal of International Business Studies*, 25 (3), p. 467-491.
- Boddewyn, J. J., et al. (1986) Service multinationals: conceptualization, measurement and theory. *Journal of International Business Studies*, 17 (3), p. 41-57.
- Boella, M. J. and Goss-Turner, S. (2005) Human resource management in the hospitality industry: an introductory guide. Oxford, Elsevier Butterworh-Heinemann.
- Bohdanowicz, P. and Zientara, P. (2009) Hotel companies' contribution to improving the quality of life of local communities and the well-being of their employees. *Tourism and Hospitality Research* 9(2), p. 147-158.
- Bohle, P., et al. (2004) Working hours, work-life conflict and health in precarious and "permanent" employment. *Revista de Salude Publica*, 38 p. 19-25.
- Bowen, J. and Ford, R. C. (2004) What experts say about managing hospitality service delivery systems. *International Journal of Contemporary Hospitality Management*, 16 (7), p. 394-401.
- Boxall, P. (1995) Building the theory of comparative HRM. Human Resource Management Journal, 5 (5), p. 5-17.
- Brand, V. (2009) Empirical business ethics research and paradigm analysis. *Journal* of Business Ethics, 86 (4), p. 429-449.
- Brien, A. (2010) Managing contingent labour in the hotel Industry by developing organisational social capital. *Journal of Human Resources in Hospitality & Tourism*, 9 (2), p. 163 - 178.
- Brookes, M., et al. (2011) Measuring competing explanations of human resource management practices through the Cranet survey: cultural versus institutional explanations. *Human Resource Management Review*, 21 (1), p. 68-79.
- Brotherton, B. (1999) Towards a definitive view of the nature of hospitality and hospitality management. *International Journal of Contemporary Hospitality Management*, 11 (4), p. 165-173.
- Brotherton, B. and Adler, G. (1999) An integrative approach to enhancing customer value and corporate performance in the international hotel industry. *Hospitality Management*, 18 (261-272).
- Burrell, J., et al. (1997) Equal opportunities for women employees in the hospitality industry: a comparison between France, Italy, Spain and the UK. *International Journal of Hospitality Management*, 16 (2), p. 161-179.
- Cairneross, G. and Kelly, S. (2008) Human resource development and 'casualisation' in hotels and resorts in eastern Australia: getting the best to the customer? *Journal of Management & Organization*, 14 (4), p. 367-384.
- Chang, Y. Y., et al. (2007) HRM strategies and MNCs from emerging economies in the UK. *European Business Review*, 19 (5), p. 404-419.
- Cho, S., et al. (2006) Measuring the impact of human resource management practices on hospitality firms' performances. *International Journal of Hospitality Management*, 25 (2), p. 262-277.
- Choy, D. (1995) The quality of tourism employment. *Tourism Management*, 16 (2), p. 129-137.

- Contractor, F. J., et al. (2010) Reconceptualizing the firm in a world of outsourcing and offshoring: the organizational and geographical relocation of high-value company functions. *Journal of Management Studies*, 47 (8), p. 1417-1433.
- Contractor, F. J. and Kundu, S. K. (1998) Modal choice in a world of alliances: analyzing organizational forms in the international hotel sector. *Journal of International Business Studies*, 29 (2), p. 325-357.
- Contrepois, S. and Jefferys, S. (2010) European Works Councils in Central and Eastern Europe: varieties of institution building among French service sector multinationals. *Industrial Relations Journal*, 41 (6), p. 584-602.
- Cooper, C., et al. (2006) New and emerging markets. *In*: Buhalis, D. and Costa, C. (eds.) *Tourism business frontiers: consumers, products and industry*. Oxford, Elsevier Butterworth-Heinemann, p. 19-29.
- Corkin, L., et al. (2008) *China's role in the development of Africa's infrastructure*. [online]. Centre for Chinese Studies, Stellenbosch University. Available from: http://www.sais-jhu.edu/bin/e/j/CorkinetalWorkingPaper.pdf [Accessed 25 / 02 / 2011].
- Costa, J. (2004) The Portuguese tourist sector: key challenges for human resources management. *International Journal of Contemporary Hospitality Management*, 16 (7), p. 402-407.
- Curtis, S. and Lucas, R. (2001) A coincidence of needs? Employers and full-time students. *Employee Relations*, 23 (1), p. 38-54.
- Cyr, D. J. and Schneider, S. C. (1996) Implications for learning: Human Resource Management in East-West joint ventures. *Organization Studies*, 17 (2), p. 207-226.
- Davidson, M., et al. (2006) Employment, flexibility and labour market practices of domestic and MNC chain luxury hotels in Australia: Where has accountability gone? *International Journal of Hospitality Management*, 25 (2), p. 193-210.
- Davidson, M. C., et al. (2010) How much does labour turnover cost? A case study of Australian four- and five-star hotels. *International Journal of Contemporary Hospitality Management*, 22 (4), p. 451-466.
- Davidson, M. C. G., et al. (2011) Hospitality HRM: past, present and the future. International Journal of Contemporary Hospitality Management Volume, 23 (4), p. 498-516.
- Deery, M. and Jago, L. K. (2002) The core and the periphery: an examination of the flexible workforce model in the hotel industry. *International Journal of Hospitality Management*, 21 (4), p. 339-351.
- Deery, M. A. and Shaw, R. N. (1999) An investigation of the relationship between employee turnover and organizational culture. *Journal of Hospitality & Tourism Research*, 23 (4), p. 387-400.
- Dev, C. S. and Olsen, M. D. (1989) Applying role theory in developing a framework for the management of customer interactions in hospitality businesses. *International Journal of Hospitality Management*, 8 (1), p. 19-33.
- Dickmann, M., et al. (2009) Exploring standardisation and knowledge networking processes in transnational human resource management. *Personnel Review*, 38 (1), p. 5-25.

- DiMaggio, P. J. and Powell, W. W. (1983) The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48 (2), p. 147-160
- Dimou, I., et al. (2003) Corporate development of international hotel firms: a transaction cost economics and agency theory approach. *International Journal of Tourism Research*, 5 (2), p. 147-149.
- Drumm, H. J. (1994) Theoretical and ethical foundations of human resource management: a German point of view. *Employee Relations*, 16 (4), p. pp.35-47.
- Dubéa, L. and Renaghanb, L. M. (1999) A look at the industry's overall best-practice champions. *The Cornell Hotel and Restaurant Administration Quarterly*, 40 (6), p. 16-26.
- Edwards, T. (1998) Multinationals, labour management and the process of reverse diffusion: A case study. *The International Journal of Human Resource Management*, 9 (4), p. 696 709.
- Edwards, T., et al. (2007) Conceptual approaches to the transfer of employment practices in multinational companies: an integrated approach. *Human Resource Management Journal*, 17 (3), p. 201-217.
- Endo, K. (2006) Foreign direct investment in tourism flows and volumes. *Tourism Management*, 27 (4), p. 600-614.
- Enz, C. A. (2009) Human Resource Management: a troubling issue for the global hotel industry *Cornell Hospitality Quarterly*, 50 (4), p. 578-583.
- Ferner, A. (1997) Country of origin effects and HRM in multinational companies. Human Resource Management Journal, 7 (1), p. 19-37.
- Ferner, A., et al. (2005) Institutional theory and the cross-national transfer of employment policy: the case of 'workforce diversity' in US multinationals. *Journal of International Business Studies*, 36 (3), p. 304-321.
- Ferner, A. and Edwards, P. (1995) Power and the diffusion of organizational change within multinational enterprises. *European Journal of Industrial Relations*, 1 (2), p. 229-257.
- Ford, R. C. and Heaton, C. P. (2001) Lessons from hospitality that can serve anyone. *Organizational Dynamics*, 30 (1), p. 30-47.
- Fournier, V. (2000) Boundary work and the (un-) making of the professions. *In*: Malin, N. (ed.) *Professionalism, boundaries and the workplace*. London, Routledge, p. 67-86.
- Gannon, J., et al. (2010) The impact of hotel management contracting on IHRM practices: understanding the bricks and brains split. *International Journal of Contemporary Hospitality Management*, 22 (5), p. 638-658.
- Geringer, J. M., et al. (2002) In search of "best practices" in international human resource management: research design and methodology. *Human Resource Management*, 41 (1), p. 5-30.
- Goeldner, C. and Ritchie, B. (2006) *Tourism: principles, practices, philosophies*. New York, Wiley
- Gonzalez, R., et al. (2011) What do we know about outsourcing in hotels? *The Service Industries Journal*, 31 (10), p. 1669 1682.

- Graf, N. S. (2009) Stock market reactions to entry mode choices of multinational hotel firms. *International Journal of Hospitality Management*, 28 (2), p. 236-244.
- Graham, M. and Lennon, J. (2002) The dilemma of operating a strategic approach to human resource management in the Scottish visitor attraction sector. *International Journal of Contemporary Hospitality Management*, 14 (5), p. 213-220.
- Guerrier, Y. and Deery, M. (1998) Research in hospitality human resource management and organizational behaviour. *International Journal of Hospitality Management*, 17 (2), p. 145-160.
- Hall, P. A. and Thelen, K. (2009) Institutional change in varieties of capitalism. Socio-Economic Review, 7 (1), p. 7-34.
- Harrington, D. and Akehurst, G. (1996) Service quality and business performance in the UK hotel industry. *International Journal of Hospitality Management*, 15 (3), p. 283-298.
- Haveman, H. A. (1993) Follow the leader: mimetic isomorphism and entry into new markets. *Administrative Science Quarterly*, 38 (4), p. 593-627.
- Haynes, P. and Fryer, G. (2000) Human resources, service quality and performance: a case study. *International Journal of Contemporary Hospitality Management*, 12 (4), p. 240-248.
- Heart, C. (1988) The power of unconditional service guarantees. *Harvard Business Review*, (7/8), p. 55-61.
- Hejazi, W. (2007) Reconsidering the concentration of US MNE activity: Is it global, regional or national? *Management International Review*, 47 (1), p. 5-27.
- Higgins-Desbiolles, F. (2006) More than an "industry": The forgotten power of tourism as a social force. *Tourism Management*, 27 (6), p. 1192-1208.
- Hjalager, A.-M. and Andersen, S. (2001) Tourism employment: contingent work or professional career? *Employee Relations*, 23 (2), p. 115-129.
- Holtbrügge, D. and Mohr, A. (2011) Subsidiary interdependencies and international human resource management practices in German MNCs. *Management International Review*, 51 (1), p. 93-115.
- Hui, O. Y. and Hsin-Wei, F. (2009) Contemporary human resource management issues and concerns in the hotel industry: identifying the priorities. *International Journal of Organizational Innovation*, 2 (1), p. 201-224.
- ILO (2001) Report for discussion at the tripartite meeting on human resources development, employment and globalization in the hotel, catering and tourism sector. [online]. International Labour Office, Geneva. Available from:

http://www.ilo.org/public/english/dialogue/sector/techmeet/tmhct01/tmhct-r.pdf [Accessed 26th November 2010].

- Ingram, H. (1999) Hospitality: a framework for a millennial review. *International Journal of Contemporary Hospitality Management*, 11 (4), p. 140-148.
- Iverson, R. D. and Deery, M. (1997) Turnover culture in the hospitality industry. Human Resource Management Journal, 7 (4), p. 71-82.
- Jackson, S. E. and Schuler, R. S. (1995) Understanding human resource management in the context of organizations and their environments. *Annual Review of Psychology*, 46 (1), p. 237-264.

- Jindra, B., et al. (2009) Subsidiary roles, vertical linkages and economic development: lessons from transition economies. *Journal of World Business*, 44 (2), p. 167-179.
- Johnson, C. and Vanetti, M. (2005) Locational strategies of international hotel chains. Annals of Tourism Research, 32 (4), p. 1077-1099.
- Jolliffe, L. and Farnsworth, R. (2003) Seasonality in tourism employment: human resource challenges. *International Journal of Contemporary Hospitality Management*, 15 (6), p. 312-316.
- Jones, T. and Sasser, W. (1995) Why satisfied consumers defect. *Harvard Business Review*, 73 (6), p. 88-99.
- Kelliher, C. and Johnson, K. (1987) Personnel management in hotels some empirical observations. *International Journal of Hospitality Management*, 6 (2), p. 103-108.
- Kelliher, C. and Johnson, K. (1997) Personnel management in hotels an update: a move to human resource management. *Progress in Tourism and Hospitality Research*, 3 p. 321-331.
- Kelliher, C. and Riley, M. (2002) Making functional flexibility stick: an assessment of the outcomes for stakeholders. *International Journal of Contemporary Hospitality Management*, 14 (5), p. 237-242.
- King, C. A. (1995) What is hospitality? International Journal of Hospitality Management, 14 (3-4), p. 219-234.
- Knox, A. (2010) 'Lost in translation': an analysis of temporary work agency employment in hotels. *Work, Employment & Society*, 24 (3), p. 449-467.
- Kostova, T. (1999) Transnational transfer of strategic organizational practices: a contextual perspective. *The Academy of Management Review*, 24 (2), p. 308-324.
- Kostova, T. and Roth, K. (2002) Adoption of an organizational practice by subsidiaries of multinational corporations: institutional and relational effects. *The Academy of Management Journal*, 45 (1), p. 215-233.
- Kusluvan, S. and Karamustafa, K. (2001) Multinational hotel development in developing countries: an exploratory analysis of critical policy issues. *International Journal of Tourism Research*, 3 (3), p. 179-197.
- Kusluvan, S. and Kusluvan, Z. (2000) Perceptions and attitudes of undergraduate tourism students towards working in the tourism industry in Turkey. *Tourism Management*, 21 p. 251-269.
- Ladhari, R. (2009) Service quality, emotional satisfaction, and behavioural intentions: a study in the hotel industry. *Managing Service Quality*, 19 (3), p. 308 - 331.
- Lai, P.-C., et al. (2008) Distancing flexibility in the hotel industry: the role of employment agencies as labour suppliers, *The International Journal of Human Resource Management*, 19 (1), p. 132-152.
- Lamminmaki, D. (2011) An examination of factors motivating hotel outsourcing. International Journal of Hospitality Management, In Press, Corrected Proof.
- Lashley, C. (1998) Matching the management of human resources to service operations. . *International Journal of Contemporary Hospitality Management*, 10 (24), p. 24-33.
- Lashley, C. (2008) Studying hospitality: insights from social sciences. *Scandinavian Journal of Hospitality & Tourism*, 8 (1), p. 69-84.

- Lau, C.-M. and Ngo, H.-Y. (1996) One country many cultures: organizational cultures of firms of different country origins. *International Business Review*, 5 (5), p. 469-486.
- Lau, C.-M. and Ngo, H.-Y. (2001) Organization development and firm performance: a comparison of multinational and local firms. *Journal of International Business Studies*, 32 (1), p. 95-114.
- Lee-Ross, D. and Pryce, J. (2010) *Human resources and tourism: skills, culture and industry.* Bristol, Channel View Publications.
- Lee, C. and Way, K. (2010) Individual employment characteristics of hotel employees that play a role in employee satisfaction and work retention. *International Journal of Hospitality Management*, 29 (3), p. 344-353.
- Lewis, C. (2008) *How the East was won: the impact of multinational companies in Eastern Europe and the former Soviet Union (1989-2004)*. London, Coptic.
- Lim, D. H. and Wentling, R. M. (1998) Modification of transferred training programmes for multinational chain hotels in Korea. *International Journal of Training and Development*, 2 (1), p. 17-27.
- Losekoot, E., et al. (2001) Conceptualising and operationalising the research interface between facilities management and hospitality management. *Facilities*, 19 (7/8), p. 296-303.
- Lowell, B. (1999) *Race for the world : strategies to build a great global firm.* Boston, Harvard Business School Press.
- Lu, Z. and Chiang, D. (2003) Strategic issues faced by Ontario hotels. *International Journal of Contemporary Hospitality Management*, 15 (6), p. 343-345.
- Lucas, R. (2002) Fragments of HRM in hospitality? Evidence from the 1998 workplace employee relations survey. *International Journal of Contemporary Hospitality Management*, 14 (5), p. 207-212.
- Lucas, R. and Deery, M. (2004) Significant developments and emerging issues in human resource management. *International Journal of Hospitality Management*, 23 (5), p. 459-472.
- Magnini, V. P. and Honeycutt, E. D. (2003) Learning orientation and the hotel expatriate manager experience. *International Journal of Hospitality Management*, 22 (3), p. 267.
- Maher, A. (1993) Measuring human resource value: an analysis based on the hotel industry. *Journal of Human Resource Costing & Accounting*, 1 (1), p. 15-33.
- Marchante, A. J., et al. (2006) Determinants of skills shortages and hard-to-fill vacancies in the hospitality sector. *Tourism Management*, 27 (5), p. 791-802.
- Martin, G. and Beaumont, P. (1998) Diffusing 'best practice' in multinational firms: prospects, practice and contestation. *The International Journal of Human Resource Management*, 9 (4), p. 671 695.
- Mathews, V. E. (2000) Competition in the international hotel industry. *International Journal of Contemporary Hospitality Management*, 12 (2), p. 114-118.
- Matias, Á., et al. (2009) An estimation of tourism dependence in French rural areas. *Advances in Tourism Economics*, p. 273-294.
- Maxwell, G. and Lyle, G. (2002) Strategic HRM and business performance in the Hilton Group. *International Journal of Contemporary Hospitality Management*, 14 (5), p. 251-252.

- Maxwell, G., et al. (2004) Quality service in the international hotel sector: A catalyst for strategic human resource development? *Journal of European Industrial Training*, 28 (2/3/4), p. 159-182.
- Mayrhofer, W. and Brewster, C. (1996) In praise of ethnocentricity: expatriate policies in European multinationals. *International Executive*, 38 (6), p. 749-778.
- McDowell, L., et al. (2007) Division, segmentation, and interpellation: the embodied labors of migrant workers in a Greater London hotel. *Economic Geography*, 83 (1), p. 1-25.
- McEvoy, G. (1984) Small business personnel practices. *Journal of Small Business Management*, 23 (4), p. 1-8.
- McPhail, R. and Fisher, R. (2008) It's more than wages: analysis of the impact of internal labour markets on the quality of jobs. *International Journal of Human Resource Management*, 19 (3), p. 461-472.
- McSherry, B. (2006) The political economy of oil in Equatorial Guinea. *African Studies Quarterly*, 8 (3), p. 23-45.
- Minbaeva, D., et al. (2003) MNC knowledge transfer, subsidiary absorptive capacity, and HRM. *Journal of International Business Studies*, 34 (6), p. 586-599.
- Mohrman, S. A. and Worley, C. G. (2009) Dealing with rough times: a capabilities development approach to surviving and thriving. *Human Resource Management*, 48 (3), p. 433-445.
- Naama, A., et al. (2008) Human resource development issues for the hotel sector in Lybia: a government perspective. *International Journal of Tourism Research*, 10 (5), p. 481-492.
- Nankervis, A. R. (2000) Human Resource Management strategies as competitive advantage: a case example of the hospitality sector in Southeast Asia & the Pacific Rim. *Research and Practice in Human Resource Management*, 8 (1), p. 111-133.
- Ondrack, D. (1985) International transfers of managers in North American and European MNEs. *Journal of International Business Studies*, 16 (3), p. 1-19.
- Orwell, G. (1933) Down and out in Paris and London. London, Penguin.
- Paik, Y., et al. (2011) Interaction effects of globalization and institutional forces on international HRM practice: Illuminating the convergence-divergence debate. *Thunderbird International Business Review*, 53 (5), p. 647-659.
- Pender, L. (2005) Managing the tourism system. *In*: Pender, L. and Sharpley, R. (eds.) *The management of tourism*. London, Sage, p. 1-13.
- Pine, R. and Qi, P. (2004) Barriers to hotel chain development in China. *International Journal of Contemporary Hospitality Management*, 16 (1), p. 37-44.
- Piso, A. (1999) Hotel and catering workers: class and unionisation. *Employee Relations*, 21 (2), p. 176-189.
- Price, L. (1994) Poor personnel practice in the hotel and catering industry: does it matter? *Human Resource Management Journal*, 4 (4), p. 44-62.
- Pudelko, M. and Harzing, A.-W. (2007) Country-of-origin, localization, or dominance effect? An empirical investigation of HRM practices in foreign subsidiaries. *Human Resource Management*, 46 (4), p. 535-559.
- Radolovic, O. (2010) Contractual liability of the guest for damage in direct hotelkeeper's contract. *Economic Research*, 23 (1), p. 97-114.

- Redman, T. and Mathews, B. P. (1998) Service quality and human resource management: a review and research agenda. *Personnel Review*, 27 (1), p. 57-77.
- Reisinger, Y. (2001) Unique characteristics of tourism, hospitality, and leisure services. In: Kandampully, J., et al. (eds.) Service quality management in hospitality, tourism, and leisure. Binghamton, Haworth Press, p. 15-50.
- Riley, M. (1996) *Human resource management in the hospitality and tourism industry.* Oxford, Butterworth-Heinemann.
- Ritzer, G. (1998) The McDonaldization thesis: explorations and extensions. London, Sage.
- Robson, P., et al. (1999) Low pay, labour market institutions, gender and part-time work: cross-national comparisons. *European Journal of Industrial Relations*, 5 (2), p. 187-207.
- Rosenzweig, P. (2006) The dual logics behind international human resource management: pressures for global integration and local responsiveness. *In*: Björkman, I. and Stahl, G. (eds.) *Handbook of Research in International Human Resource Management*. Cheltenham, Edward Elgar, p.
- Rosenzweig, P. M. and Nohria, N. (1994) Influences on human resource management practices in multinational corporations. *Journal of International Business Studies*, 25 (2), p. 229-251.
- Rosenzweig, P. M. and Singh, J. V. (1991) Organizational environments and the multinational enterprise. *The Academy of Management Review*, 16 (2), p. 340-361.
- Rowden, R. W. (1995) The role of human resource development in successful small to mid-sized manufacturing business: a comparative case study. *Human Resource Development Quarterly*, 6 (4), p. 355-373.
- Rowley, G. and Purcell, K. (2001) 'As cooks go, she went': is labour churn inevitable? *International Journal of Hospitality Management*, 20 (2), p. 163-185.
- Rutherford, D. G. and O'Fallon, M. J. (2007) *Hotel management and operations*. Hoboken, NJ, John Wiley & Sons.
- Sartorius, K., et al. (2011) Human resource management and cultural diversity: a case study in Mozambique. *The International Journal of Human Resource Management*, 22 (9), p. 1963-1985.
- Saunders, M., et al. (2007) *Research methods for business students*. Harlow, Pearson Education UK.
- Schuler, R. S. and Rogovsky, N. (1998) Understanding compensation practice variations across firms: the impact of national culture. *Journal of International Business Studies*, 29 (1), p. 159-177.
- Schuler, R. S. and Tarique, I. (2007) International human resource management: a North American perspective, a thematic update and suggestions for future research. *The International Journal of Human Resource Management*, 18 (5), p. 717 - 744.
- Shamir, B. (1980) Between service and servility: role conflict in subordinate service roles. *Human Relations*, 33 (10), p. 741-756.
- Shay, J. P. and Baack, S. A. (2004) Expatriate assignment, adjustment and effectiveness: an empirical examination of the big picture. *Journal of International Business Studies*, 35 (216-232).

- Simons, T. and Enz, C. A. (1995) Motivating hotel employees. *Cornell Hotel and Restaurant Administration Quarterly*, 36 (1), p. 20-27.
- Sisson, K., et al. (2003) All benchmarkers now? Benchmarking and the 'Europeanisation' of industrial relations. *Industrial Relations Journal*, 34 (1), p. 15-31.
- Soltani, E. and Wilkinson, A. (2010) What is happening to flexible workers in the supply chain partnerships between hotel housekeeping departments and their partner employment agencies? *International Journal of Hospitality Management*, 29 (1), p. 108-119.
- Szulanski, G. (1996) Exploring internal stickiness: impediments to the transfer of best practice within the firm. *Strategic Management Journal*, 17, p. 27-43.
- Tanke, M. L. (2001) *Human resource management for the hospitality industry*. Albany, Delmar.
- Taylor, S. and Edgar, D. (1996) Hospitality research: the emperor's new clothes? *International Journal of Hospitality Management*, 15 (3), p. 211-227.
- Teare, R. (1995) The international hospitality business: a thematic perspective. International Journal of Contemporary Hospitality Management, 7 (7), p. 55 -73.
- Timo, N. and Davidson, M. (2005) A survey of employee relations practices and demographics of MNC chain and domestic luxury hotels in Australia. *Employee Relations*, 27 (2), p. 175-192.
- UNCTAD (2004) World investment report 2004: the shift towards services. [online]. Available from: http://www.unctad.org/Templates/WebFlyer.asp?intItemID=3235&lang=1 [Accessed 17th September 2009].
- Walsh, T. J. (1990) Flexible labour utilisation in the private service sector. Work, Employment & Society, 4 (4), p. 517-530.
- Whitla, P., et al. (2007) Global strategies in the international hotel industry. International Journal of Hospitality Management, 26 (4), p. 777-792.
- Whitley, R. (1999) Divergent capitalisms: the social structuring and change of business systems. Oxford, Oxford University Press.
- Wilton, N. (2006) Strategic choice and organisational context in HRM in the UK hotel sector. *Service Industries Journal*, 26 (8), p. 903-919.
- Wood, G. (2004) Business and politics in a criminal state: the case of Equatorial Guinea. *African Affairs*, 103 (413), p. 547-567.
- Wood, R. C. (1992) Working in hotels and catering. London, Routledge.
- Woods, R. H. (1999) Predicting is difficult, especially about the future: human resources in the new millennium. *International Journal of Hospitality Management*, 18 (4), p. 443-456.
- Woodworth, R. M. and Mandelbaum, R. (2010) Seventy-five years of U.S. hotel revenues, expenses, and profits. *Cornell Hospitality Quarterly*, 51 (1), p. 20-26.
- Wright, P. M., et al. (2005) New models of strategic HRM in a global context. *The International Journal of Human Resource Management*, 16 (6), p. 875 - 881.
- WTO (2001) Tourism 2020 Vision: global forecasts and profiles of market segments. Madrid, WTO.

- WTO (2011) International tourism 2010: multi-speed recovery. [online]. Available from: http://85.62.13.114/media/news/en/press_det.php?id=7331&idioma=E [Accessed 27/06/2011].
- WTTC (2009) *Travel & tourism economic impact: executive summary*. [online]. Available from: http://www.wttc.org/bin/pdf/original_pdf_file/exec_summary_2009.pdf [Accessed 07/07/2010].
- WTTC (2011) *World -key facts at a glance*. [online]. Available from: http://www.wttc.org/eng/Tourism_Research/Economic_Research/ [Accessed 17/02/2011].
- Yan, Y., et al. (2007) Vertical integration of corporate management in international firms: implementation of HRM and the asset specificities of firms in China. *International Journal of Human Resource Management*, 18 (5), p. 788-807.
- Yeung, A. (2006) Setting people up for success: how the Portman Ritz-Carlton hotel gets the best from its people. *Human Resource Management*, 45 (2), p. 267-275.
- Zhang, H. Q. and Wu, E. (2004) Human resources issues facing the hotel and travel industry in China. International Journal of Contemporary Hospitality Management, 16 (7), p. 424-428.
- Zuehl, J. J. and Sherwyn, D. S. (2001) Identifying uniform employment termination practices for multinational employers. *Cornell Hotel and Restaurant Administration Quarterly*, 42 (5), p. 72-85.

Appendix I

List of 7 properties fully owned by HotelCo participating to the survey

• EAME-international properties:

> CONTINENTAL EUROPE

- United Kingdom (1 hotel)
- Switzerland (1 hotel)
- Germany (3 hotels)

> CAUCASUS (area between Eastern Europe and Western Asia)

- Azerbaijan (1 hotel)
- > CENTRAL ASIA
 - Kyrgyzstan (1 hotel)

Family Business in Greece Investigation of the basic problems

Sotirios Vlachakis¹, Marja Naaranoja², and Kerstin Siakas³

¹ PhD Student in Business Administration (Department of Industrial Management), University of Vaasa, Finland

sotirisv@mkt.teithe.gr

² University of Vaasa, Department of Industrial Management, P.O. Box 700 65101 Vaasa,

Finland,

marja.naaranoja@uwasa.fi

³ Alexander Technological Educational Institute of Thessaloniki, Department of Informatics,

P.O. Box 141, 57400 Thessaloniki, Greece siaka@it.teithe.gr

Abstract. Little research has so far been carried out concerning the Greek family businesses. According to the Grant Thorton research in 2006, 80% of entrepreneurs consider their business to be a family business. The existing international bibliography shows that the main complications within a family business are as follows: a) succession b) intra-business and intra-family controversies between family members, c) the sequence in leadership and finally, d) the effects these previous issues have on the strategies a family business follows. The gap in the literature regarding Greek family businesses, does not allow us to accept that this is also true for Greek family businesses.

This paper aims to investigate the main problems faced by family businesses in Greece. Based on a primary field-study which is focused on 30 family business-es it examines whether and how family businesses have been affected by the social-economical development model of the Greek economy, as well as by the notion and culture of the Greek family. The methodology of this paper is based on in-depth interviews with leading members of each family business. The interviews were based on the expected future developments (scenarios). Survey results show that a) the problems mentioned in the international literature as the main problems of family business occur also in Greek family business, b) there are some peculiarities due to the socio-economic system, the prevailing business culture and the family business, which relate to the succession process, the unstable tax environment, the low extraversion of SMEs and the existence of forced and self-made entrepreneurship in Greece.

Keywords: Family business, entrepreneurship, SMEs

1 Introduction

Research carried out on family businesses in Greece is very limited. There are many individual reports as well as reports by independent institutions presenting survey results. These results have occasionally handled with Greek family businesses. They all conclude that the vast majority of enterprises in Greece (over 80%) are family business (Grand Thorton, 2006; Industrial Chamber of Athens, 2010; PWH, 2010).

Regarding the size of the family business, there is an assumption that the Greek family businesses are SMEs, especially small and very small firms. This assumption is based on general statistics about entrepreneurship, indicating that there are more than 770.000 businesses in Greece, employing 1-10 employees and comprising more than 98% of total Greek business (SMEs in Europe, 2008;ESEYE, 2011).

According to the existing literature on family businesses it is not clear if family businesses in Greece face the same problems as the family businesses in other countries. More over, we do not have clear information on whether and how the Greek socio-economic system, the prevailing business culture or the structure of the Greek family affect the structure, the development or the every day's operation of the Greek family businesses.

The motivation to carry out research on Greek family businesses has a dual purpose to investigate a) if family businesses in Greece face the same problems as the family business in other countries, b) if they have been affected in a special way by the social-economical development model of the Greek economy, as well as by the notion and culture of the Greek family.

In this paper, based on a primary research, we make the attempt to link the reported problems of family businesses in the literature, to those faced by family businesses in Greece.

2 Literature review

A systemic approach to family business research started in the 60's. The family business is a composition of two overlapping and interrelated subsystems. Because of the overlapping the same people, family and business members, incur obligations to both subsystems (P. Davis, 1983; Kepner, 1983). The problem is to find a strategy that simultaneously serves the needs of a) individual family members b) of "family sys-tem" c) of "business system" (Whitesite & HerzBrown, 1991;Gersick et al., 1997).

The two cycle model evolved in the '80s to a three cycle model. The new subsystem is the acquired property. Any family member owns part of the propertyownership patterns, may influence company decisions (Lansberg, 1983).

The family business system has not been examined as a "dynamic" system which is in interaction with other external systems. An answer to this reflection was given through the Open-Systems Approach. According to that the "family business" is a system composed of several subsystems and simultaneously the "family business" itself is a system that is a member of the great economic and political system (Pieper et al., 2007). The environment consists of the company's customers, competitors, suppliers, the state and other social institutions and organizations. Same researchers argue that current research omits the "human" in the study of family business. Also Zahra emphasize the "human factor as a module in family business". Everyone involved in the family business, founder entrepreneur, simple family member, future successors or "administrative staff non family member" affects every subsystem in a unique way and may simultaneously belongs to more than one subsystem. (Zahra et al., 2006)

Another key parameter which preoccupied the researchers was the "business culture". For the family business the concept of business culture is complicated. The company's founder is a primary creator of the business culture of the family business (Barney, 1986). Because of the dominant role of the founder his/her personal values and incentives are key factors of the configuration of the business culture. This does not happen only during the first years, but also when the business passes to the next generation (Denison et al., 2004). Dyer acknowledged four different types of business culture in the family business a) paternalistic, b) liberal, c) participatory and d) person-centered (Dyer, 2003).

The interaction of several sub-systems and individuals in the "family business" system can cause particular problems (Denison et al., 2004). Neubauer and Lank (1998) have grouped all the particular problems into four major categories a) strategy, b) governance, c) intra-business and intra-family controversies and d) successions problem.

A. Strategic plan in family businesses

Strategic plan is by Kefalas (2008) defined as an informal plan that exists in the founder's mind. It is not communicated to all stakeholders of the company, rarely applied methodically and even more rarely formally maintained.

According to Bhalla the strategic decisions there are only a few strictly economic and business issues taken by the family business. The family business is trying to build or its strategic decisions a unique business culture, based on a common vision of all members (Bhalla et al., 2006). The personal character of the administration and the emotional involvement of the owners in strategic planning maybe have a negative affect by the implementation of targeted strategies, affecting the financial figures of the company (company's profit in relation to the competition, sales figures, etc.). (Koufopoulos et al., 2011) For an understanding of strategic planning in family businesses, Pieper and Klein, suggest a model of four theoretical approaches. The purpose of this model is to understand the progressing process of the strategy in family businesses and what is produced as a result of it (Pieper & Klein, 2007). The four approaches are:

1) Moral/ethical values. The basic moral values of the family assimilated by the younger generations diffuse in family business and form the basis for decision making. Strategic decisions and tactics tend to be heavily influenced by the fundamental family values.

- 2) Roles. The interaction of the different roles of each member in a family business is a key feature. If the roles are not clear people find it difficult to work constructively and to participate in a unified strategy for the business.
- 3) Arena (the place of the meetings and decisions). Strategic decisions but also less formal and important decisions in family firms are made not only in professional business offices but they also are extended at home.
- Rationality/logicality. The interaction of the systems "family" and "business" and the simultaneous existence of strong emotional ties infuse the decision making process with more emotion than logic. Concluding we pose our Hypothesis:

Strategic plan is an informal plan which is based on the funder's vision; nonetheless strategic decisions and tactics tend to be heavily influenced by the fundamental family values, strong emotional ties infuse the decision making process with more emotion than logic.

B. Governance

Suare and Santana-Martin define governance in family business as "capitularies and mechanisms aimed to determine the relationships within the family and the relationships between the "family" and the "business". The capitularies and mechanisms formulate specific roles for each member of the family associated with governance, enhance the effectiveness of the management team, stimulate the same vision of the family and effectively assist in resolving any disputes (Suare & Santana-Martin, 2004). In family businesses, there is a large turnout of family members in the governance process. Creating an informal family council is very common. It is very useful especially a) when the family business is still in its first generation, b) future successors have an applicable age and c) future successors begin to show interest in the company (Neubauer & Lank, 1998). Through "informal family councils" messaging to the next generation are communicated regarding the vision and values of the family business as well as methods of administration (Ward, 2010). According to Kets de Vries et. al., (1997) the family council should be a continuous and more organized process. Existing research shows that many family businesses are a source of innovation at the beginning of their life cycle but later on most of them show reluctance and fear of change and follow conservative strategies (Shepherd & Zahra, 2003). Possible reasons for this is that some owners maintain control for a very long time leaving no room for administrative renewal, the fear of possible failure and the close relationship of the family property with the family business (Zahra, 2005).

Concluding we can pose the Hypothesis: There is a large turnout of family members in the governance process. Governance is effective when members have specific roles based on commonly accepted capitularies and mechanisms. Creating an informal family council is very common.

C. Conflicts

Existing literature refers, that within the family business there are many conflicts, mainly due to the fact that there are two interlocked different systems, "the family" and "the business" The characteristics of these two systems are in conflict with each other. Based on the dipole "Business-Family", inside of the family business lurks a default conflict (Kefalas, 2008).

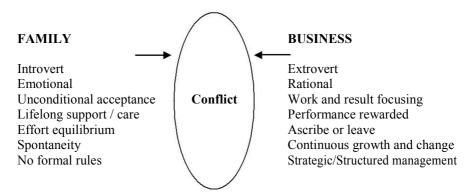


Figure 1. The conflicting roles of family and business (based on Kefalas, 2008)

The family operates according to some rules, which can be formal or informal involving sentimentality. Intense emotion in the family can bring many negative effects on business (Levinson Harry, 1971). The lack of organizational structure in management is also a source of conflicts because as a result the roles among the members of the family business are not clear (Grote Jim, 2003). The causes of conflict in family businesses can be different from company to company. We first look at the conflict areas, where the main causes occur. These areas are:

- 1) Control. Most conflicts in a family business concern issues of control and management of the company (Grote Jim, 2003).
- 2) Business goals against family's objectives (Bedeian et al., 1988).
- 3) Preparations of the succession plan (Grote, 2003).
- 4) Wage levels (Bedeian et al., 1988).
- 5) Levels of ownership (Levinson, 1971).
- 6) Acquisitions of property between members (Bedeian, 1988).
- 7) Future business strategy (Grote, 2003).
- 8) Equal responsibilities and duties but unequal offer (Hubler, 1999)
- 9) No right of dismissal, regardless if someone gives / offers or not (Kefalas,2008),
- 10) Different work style and different Lifestyle among members (Levinson, 1971).
- 11) The role of relatives by marriage (Bedeian et al., 1988).

Conflicts exist in all these combinations because there is more emotion and less business logic involved. Researchers suggested occasionally some ways to prevent or manage the conflicts. (Grote Jim, 2003). Suggestions identified in the literature are:

- 1) Separating the problem from the person.
- 2) Positive attitude towards things.
- 3) Mutual respect between the members and the existence of the habit of discussing problems and allowing every member to express his/her opinion openly.
- 4) Compromise in making key decisions.
- 5) Encouragement and creation of written and oral communication.
- 6) Statutory family gatherings or councils at specified intervals or on specific important issues.
- 7) Written management code.

Concluding we can pose the Hypothesis:

In family business there exist conflicts between family members due to the coexistence and interaction of the two sub-systems "family" and "business". Conflicts exist in several combinations because there is more emotion and less business logic involved.

D. The successions process

Family businesses worldwide have large failure rates of transfer to future generations. One of the most important factors that can influence the future and longevity of a family business is the process of selecting a successor. Beckhard & Dyer (1983) found that 70% of all family businesses are sold or liquidated after the death or retirement of their founders. In total Gimeno et al (2010) similarly found that 70% of family firms cease their activity during the transition from the first to the second generation and only 15% of the family businesses survive until the third generation. The younger family members who decide, among other career options, to join the family business are driven by a sense of need or desire which is called "commitment to the family business» (Sharma & Irving, 2005).

According to the researchers we find four different levels or types of family business successor's commitment, namely:

- 1) Affective commitment (wish). The successor with affective commitment accepts fully and feels loyalty and enthusiasm about the business objectives (Meyer & Allen, 1991)
- 2) Normative commitment (must). The commitment is driven by a sense of duty and obligation towards family and business (Stewart, 2003).
- Calculative commitment (profit). The commitment is stimulated by the economic and social rewards offered by the company (Sharma & Rao, 2000; Sharma & Irving, 2005).
- 4) Imperative commitment (need). The successor believes that he/she will not be able to work efficiently outside of the family business and therefore he/she feels that his/her career in the family business is the only realistic option (Cannela& Shen, 2001).

Many researchers believe that a succession design itself is not enough to ensure a successful succession in a family business. The preparation of succession should be a continual and lengthy process. The goal is not only the short-term successful transition to the next generation. Future functioning is also an important goal (Astrachan, 2001; Murray, 2003). All social, economic, legal, ethical and strategic dimensions, that influence the succession process, should be taken into account over time (Sten, 2001). Lambrecht and Donckels (2006), created the ideal path which must be followed over time for a successful succession. They consider that key factors are the business, the family and each family member separately as a person actively involved in the business or not. These three entities should interact, influencing each other in contributing to strategy and shaping moral values. The first step requires that the younger members of the family must have contact with the business. The second step should ensure the best possible education of the successor by academic education, experiential training (on-the job) and training in other companies. Finally, in the third step the formal involvement of younger members of the family is done gradually, beginning from the lowest possible position in the family business (Lambrecht & Donckels, 2006).

A similar approach to successful succession in steps was referred by Kefalas (2008), who considers that succession is basically an issue of individual choice of the founder of the family business. The steps proposed by Kefalas for a successful succession process are three: Step one "Values transfer". Building successor's personal character and clarifying the principles of business. Second step "Management transfer". Third step "Property transfer". Kefalas conclude that the succession to the next generation is a long complicated process and not just an event (Kefalas, 2008).

Concluding we can pose the Hypothesis:

Family businesses worldwide have large failure rates of transfer to future generations. The most important factor is the process of selecting a successor. It must be designed and prepared in steps, in a continual and lengthy process.

3 Research methodology

The study reported in this paper used a qualitative research methodology comprising in-depth interviews with leading members of each family business. The interviews were based on the expected future developments (scenarios) of the family business. The target group consisted of Greek medium and small family businesses, where the total number of employees is less than 50. The interviews were conducted in 30 family businesses of Northwest Thessaloniki, from 01/10/2011 until 31/03/2012. The characteristics of the family businesses interviewed are:

- 1. the majority of the adult family members are employed in the family business
- 2. the management of the company is practiced by the members of the family.
- 3. the income obtained from the family business is the main financial source for the family.
- 4. the members of the family have inherited the business from previous generations or in case they are the founders themselves, they have explicitly stated their intention to transfer the business to future generations.
- 5. the business is running in the Prefecture of Thessaloniki.

The method of gathering the information was based on personal in-dept interviews with leading members. The interviews were structured and based on a core of questions defined and justified after a pre-testing stage including interviews of three family businesses. The interviews were based on expected future developments (scenarios). The scenarios were chosen based on two criteria, namely a) the literature review

b) expected future developments.

This methodology (personal in-depth interviews based on the above scenarios) is chosen as Wilson (2006) puts it to achieve higher acceptance from the interviewee, easier explanation of fuzzy questions, and finally possibility of further analysis of the scenarios.

4 Empirical analysis

Findings

Strategic plan

H: 1 Strategic plan is an informal plan which is based on the funder's vision; nonetheless strategic decisions and tactics tend to be heavily influenced by the fundamental family values, strong emotional ties infuse the decision making process with more emotion than logic.

Strategic plan, an unknown concept

The strategic plan is a concept, unknown to the vast majority of Greek smallmedium family businesses. It does not exist neither as a written instrument for operation and development, nor as a tool for sharing responsibilities among the members of the family business. Usually, there are only individual goals, which are determined by current market situation (eg. sales targets, new product(s), new customers etc.). The goals are usually short-term and they are strongly influenced by the general climate of the economic crisis, which include lack of information and alternative solutions.

The strategic plan exists in the mind of the manager, who is usually the founder of the family business. He/she discusses it in general with the main company's members (spouse, older children). The discussion concerns mainly the question: what are we doing now? Generally, there is a lack of methodical application of measures, based on a specific plan. As the main causes of this situation we can mention:

- a) The owner's involvement with the everyday problems. He/she spends time dealing with everyday problems. The future design is delayed as a consequence of that the fact that today's problems are considered to be more important than the future ones.
- b) The lack of knowledge and fear of failure. The small-medium business owners do not have the necessary knowledge to develop a medium term or a long term plan. After the creation of wealth, they are fear of firing at risk the total family achievement.

Some important decisions are often taken under the influence of emotional charging.

This study identified some reasons for emotional charging in strategic decisions, as follows:

- a) when a decision has to be taken about the participation of a family member in the business, the future prospective and evolution of their children in the family business is taken seriously into consideration. If the family business is struggling for survival the parents may consider that a career outside the family business will be a better choice for the child. In general Greek families are very protective towards their children and most of the parents engage to a very high degree into the future career choices of their children.
- b) when, in difficult financial situations, they are forced to mortgage personal assets of family members, especially when they predate the founding of the company,
- c) when the enterprise's assets or the prosperity and reputation of the family could be staked.

Governance

H: 2 There is a large turnout of family members in the governance process. Governance is effective when members have specific roles based on commonly accepted capitularies and mechanisms. Creating an informal family council is very common.

Family council, an informal tool for decision making

In this study we found that the strategic or important decisions are taken, at a rate greater than 60%, from the company's founder. He/she is supposed to understand the challenges better than anyone else in the business The founder also is more aware of the future goals of the family business than anybody else. In most of the cases the founder discusses the problems with the main company members, but he/she always makes the final decision. In the remaining 40% of companies there exists a family council. Its role is advisory. It can be characterized as an informal tool, because it does not operate in accordance with predetermined procedures, but usually spontaneously. It contributes more to coil all the family business members around the common goal. In very few cases (7%) and only in larger business, the final decisions are taken based on the majority principle. There are also very few cases where the final decisions are taken by non family members.

There is a continuous interaction between the subsystems "business" and "family". The problems of the company are transferred to the family and vice versa. This is because the same people have dual roles in the family and in the business. It is very difficult to make a clear distinction. The testimonies of the interviewees are very characteristic. "When we gather on Sundays for lunch the two brothers (the two main shareholders) continuously talk about the business". (Spouse not working in the business). "At work I can chat with my kids about all kind of subjects. When I go home I'm very tired and I have to do the housework" (Mother working in family business). "Things at work are running very fast. Sometimes we have no time to think about the work" (Woman, owner of a family business). "I see my mother more at work. In my spare time I want to be with my friends" (Daughter working in family business). "At home in the evening I have the ability to calm down and put things in order" (Man directing a family business, second generation). "In the office everyone is expecting something from me and the phone rings all the time. At home I have the

peace to think better" (Man managing a family business).

The workplace tends to be the place where current daily chores are processed. All interviewees consider their daily tasks to be very important and devote most of their time for them. When business functions are clearly distinct, members decide for strategic or other important issues within the workplace.

On some issues there may be strict rules for operation. This research identified the following:

a) cash-flow management;

b) reinvestment of funds;

c) relationships with customers / suppliers;

d) relationships with the state.

The need for economic decision making rules is during the economic crisis stronger than earlier since the owners are afraid of fast liquidation

Succession

H: 3 Family businesses worldwide have large failure rates of transfer to future generations. The most important factor is the process of selecting a successor. It must be designed and prepared in steps, in a continual and lengthy process.

"We have to do something for our children"

The process of selecting and preparing a successor seems to be a very important issue for Greek family business owners. In total 21 of the businesses interviewed reported that they have adult children who are more than 65%,, employed by the family business to even periodically. Most of the Greek family businesses (more than 80%) have a succession plan. The high degree of preparation for succession is likely to be due to the uncertainty working environment in time of crises, but also to the high percentage of uncertainty avoidance index reported for Greece by Hofstede (2001). The expression all the interviewees stated is very characteristic, "we have to do some-thing for our children".

Additionally there are also some other important factors influencing the succession planning. Over the last thirty years, there have been significant changes in Greece in the sector of Higher Education (HE). Today it is easier for the younger generations gain entrance to HE, compared with their parents (currently 50-60 years old) who had very limited opportunities. Today's parents want to provide an opportunity for their children to have good education, because they believe that in such a case they will be better equipped for their future career choices.

By the question if children's studies are relevant to the company's subject two trends were distinct, namely:

- a) The larger the company, the more distinct are the business functions, leading to a higher possibility to attract new family members of the new generation to engage in the family business. Results from our study show that more than 70% of the family's children study something related to the company's subject.
- b) When the business is small, it is not taken for granted that it can ensure the future career perspective of all the children within the company. In such a case some of the children follow different studies in a proportion of more than 55%.

The business take care of the educational cost and the start up of the career of these children. It also was observed that a large percentage of the children's education (80%) is relevant to the subject of business when particular expertise is required (eg. confectionery or pharmacy) and when the company has good growth rates. By the question to the parents "witch one position do you think is good for your children to start their career in the family business, when they have finished their studies", we

had four different levels of answers, namely
a) starting from the lowest level - 34%;
b) starting with few responsibilities - 26%;

- c) managerial position with increased responsibilities 18%;
- d) managerial position with decision-making power together with the founder 22%.

Conflicts

H: 4 In family business there exist always conflicts between family members due to the coexistence and interaction of the two sub-systems "family" and "business". Conflicts exist in several combinations because there is more emotion and less business logic involved.

"It is impossible that there are no conflicts"

- a) About the existence of conflicts within the Greek family businesses, is important to mention that nine out of the ten interviewees, when they heard this question smiled and said characteristic: *"It is impossible that there are would be conflicts between the members"*. They seemed to address the conflict's issue calmly and positively, because this is a part of their daily routines. They show receptivity for the conflicts and they try to overcome them, based on the good relationship that obviously exists between the family members. *"The love between us is unconditional nobody is going to hurt another family member"*. About the factors that can cause conflict, we found that these are the following:
- a) The more obscure and ambiguous the power, job-roles and the responsibilities are, the more it is likely that there are some conflicts, due to the existence of informal multi-stakeholder governance. "Sometimes we entangle in each other's work and this leads to conflicts".
- b) The generational gap that obviously exists in all sectors nowadays is enhanced by the adaption of innovative elements, such as the use of Information Technology (IT) and marketing, which the new generation easily adopts. The first generation must first be convinced of their usefulness in order to allow for their implementation. *"He*

(the founder) became more receptive only when he saw that the problems are solved".

- c) Root cause of conflicts is the involvement of the parents in the business, although the process of succession is already completed. "Even now he wants to have the control, because he feels more useful".
- d) The level of wages is not a major cause of conflict, when all the company's members come from the same family. The company ensures that any conflicts can be settled in a "fair" way. But when the business members come from more than

one family, especially when the second generation enters the business, strong conflicts can be created. The main causes are the different levels of education and training of new members and the claim for jobs with different responsibilities.

e) When the founder applies an authoritarian management model, children cannot easily accept their father/mother as a strict employer. In these cases everyday issues can be a cause of conflicts.

Greek socio-economical business environment.

H: 5 Family businesses in Greece face particular problems due to the Greek socioeconomical business environment.

We asked all respondent family businesses if they face other significant problems. We gave, in this point, the opportunity to the interviewees to respond freely about anyone other issues which they consider to be important. We received answers about problems which do not solely relate to family businesses, but which are important for the general business climate. Because the study was carried out in a time of economic crisis, inevitably a lot of reference to the crisis were made such as:

- a) Family firms were considered to be more resilient to the economic crisis (also confirmed by Piraeus Traders Association, 2009).
- b) The massive entry of large multinational companies, which afford strong brand names in the market, was perceived as a potential risk for SMEs.

c) The massive imports of cheap products from China were perceived as a potential risk Problems that can be characterized as specific to Greek family businesses,

were reported as follow:

a) <u>The unstable tax system in Greece</u>

Although it applies to all Greek companies, in the case of family businesses it is particularly important. They are taxed during their operation in a unique triple way, as company's turnover, as individual income and as property. More pronounced are the taxation's problems by the transfer of the business to the next generation, especially regarding its assets. The managers of family businesses are confused because they are facing three different tax situations, without knowing in advance which is favourable

a) when the assets are owned by the company b) when the assets are owned by family members c) when the firm uses only rented assets, and family members create assets outside of the family business. The Greek family business owners trying to position their company to acquire its own facilities, because they believe that this is a competitive advantage, but after many changes to the tax system it can be proven that it is not favourable. The main problems arise when the company and all its assets are transferred to the second generation. In this case the number of tax law changes may incur further the family business.

b) <u>Self-created entrepreneurs</u>, necessity (push) entrepreneurship

There is an overwhelming majority (over 55%) of self-created entrepreneurs. They started their own business due to an unstable working and business environment, mainly in the decades of '70 and '80. These entrepreneurs founded their business because of the need to find a job for themselves and for their family members. It was a "necessity entrepreneurship" not an "opportunity (pull) entrepreneurship". They

usually had a low level of education and they were prepared to deal with any business subject, they believed to be profitable. In the 90's the business environment had improved significantly. Young people with good education created "opportunity (pull) entrepreneurship". Today, unfortunately, the economic environment is very unfavorable. Necessity (push) entrepreneurship is prevalent. The creation of small family businesses or the retaining of young people in existing family businesses is almost a unique employment solution for the new generation.

c) Low extraversion

- Regarding the level of extroversion in the Greek family business, we observed that:
 - 1) Innovation is absent from all enterprises of our study
 - 2) The search for new markets is a constant challenge. The tool witch usually is used is internet searching
 - 3) The export rate is limited.
 - 4) There is no knowledge of marketing tools. Marketing is limited to very basic things.
 - 5) The use of external consultants is almost nonexistent.

The reasons for the low extroversion are mainly the small size of the company, the lack of knowledge regarding innovation and the use of marketing tools, as well as of the low availability of funds.

5 Conclusions

Following the above analysis, the hypothesis, that the Greek family business faces the same problems as family businesses in other countries, is verified. These problems have been mentioned and analyzed extensively in the international literature. Our findings contribute and verify as follows:

- 1. Regarding the strategy, we found that the role of company's founder is determinant for the Greek family business. There is no a clear strategic plan in the Greek family business, which can motivate all its members. The business plan rarely works as a tool of written rules. It exists in the founder's mind, who maybe discusses it in general with the main company's members (spouse, older children), but the founder almost always takes the final decision. Many times some important decisions can also be taken under the influence of emotional charging.
- 2. As regards the administration and the daily business operation, our study shows that there are clear rules of operation only on certain issues, such as treasury management, relationship with customers and relationship with suppliers. The role of these rules was reinforced by the recent economic crisis. The family council is, for the vast majority of family businesses, an informal advisory body. The place where decisions are made can be both the work-place, and the home.
- 3. About the conflicts between the members we have found that in Greek family businesses there are similar conflicts within family businesses as nother countries. The main factors that cause these conflicts are the following order of evaluation.
 - a. The ambiguity of the distribution of power, job-roles and responsibilities of the members, which leads to multi-stakeholder management

- b. The generation gap which is enhanced by the ability of the new generation in using new technologies and marketing tools.
- c. The involvement of the first generation in company's control, even if the succession process has been completed.
- d. The wages level. This became to be more important in case there are members of more than one family in the business.
- e. When the founder applies an authoritarian management model, children cannot accept their father/mother as a strict employer.
- f. Some everyday's issues can also be a cause of conflicts.
- 4. On the issue of succession we observed that especially by the Greek family businesses, succession is a process that is systematically prepared. The education of the children is largely related to the company's subject. The children are employed by the business, even occasionally when they study. This succession's preparation is a peculiarity of the Greek family business due to the unstable work and business environment. "To create something for our children" is the main thought of first generation.
- 5. Some peculiarities of the Greek family businesses due to the socio-economic system of the country were identified, including:
 - a. The unstable tax system. This affects the family business during its operation and by the transfer to the next generation.
 - b. The founders of family firms are self-created entrepreneurs. They founded their own business to secure work for themselves and for the members of their family. They operate more under conditions of "necessity (push) entrepreneurship" and less of "opportunity (pull) entrepreneurship".
 - c. There is a low level of extraversion by Greek family businesses due to the small company's size, the lack of skills in using marketing tools, and the low availability funds.

Generally, the Greek family business faces the same problems as the family business in other countries. Moreover they face some particular problems due to socioeconomic system of the country.

Acknowledge

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research

Funding Program: ARCHIMEDES III Investing in knowledge society through the European Social Fund.

6 References

- Astrachan, J. H.; Klien, S. B., & Smyrnios, K. X. (2002). The F-PEC scale of family influence: A proposal for solving the family business definition problem. *Family Business Review*, **15**(1):45–58.
- Beckhard, R. & Dyer, G.W. Jr. (1983). Managing continuity in the family-owned business. *Organizational Dynamics*, **12**(1):5–13.
- Chua; J. H, Chrisman, J. J. & Sharma, P. (1997). Strategic Management of the Family Business: Past Research and Future Challenges, *Family Business Review*, **10**(1): 1-35.
- Chua; J. H, Chrisman, J. J. & Sharma, P. (1999). Defining the Family Business by Behavior. *Entrepreneurship Theory and Practice*, **23**(4):19-39.
- Churchill, N. C., & Hatten, K. J. (1987). Non-market based transfers of wealth and power: A research framework for family businesses. *American Journal of Small Business*, 11(3):51–64.
- Danco, L. (1975). Beyond survival: A business owner's guide for success. Cleveland: University Press.
- Danes, S. M.; Rueter, M. A.; Kwon, H.-K., & Doherty, W. (2002). Family FIRO model: An application to family businesses. *Family Business Review*, 15(1):31–43.
- Davis, P. (1983). Realizing the potential of the family business. *Organizational Dynamics*, **12**(1):47–57.
- Davis, P., & Stern, D. (1980). Adaptation, survival, and growth of the family business: An integrated systems perspective. *Human Relations*, **34**(3):207-224.
- Denison, D.; Lief, C. & Ward, J. L. (2004). Culture in Family-Owned Enterprises: Recognizing and Leveraging Unique Strengths, Family Business Review, 17(1):61-70.
- Dyer Jr., W. G. (2003). The family: The missing variable in organizational research. *Entrepreneurship Theory and Practice*, **27**(4):401–416.
- Gersick, K. E.; Davis, J. A.; McCollom, H. M., & Lansberg, I. (1997). *Generation to generation: Life cycles of the family business*. Boston, MA: Harvard Business school Press.
- Grote J. . "Conflicting Generations: A New Theory of Family Business Rivarly" . Family Business Review 16, no.2, June 2003.
- Hofstede Geert (2001): Culture's consequences: comparing values, behaviours, institutions, and organisations - 2nd Ed. - Thousand Oaks, Calif.; London: Sage Publica-tions
- Kets de Vries M. F. R. (1993), The dynamics of family controlled firms: the good and the bad news, *Organizational Dynamics*, **21**(3):59-71.
- Kets De Vries; Manfred F.R.; Carlock, Randel S. & Florent-Treacy, E. (2007). Family Business on the Couch: A Psychological Prospective. West Sussex, UK: John Wiley & Sons Ltd.
- Lansberg, I. (1983). Managing human resources in family firms: The problem of institutional overlap Organizational Dynamics, 12(1):39–48.
- Litz, R. A. (1995). The family business: Toward definitional clarity. In *Proceedings* of the Academy of Management (pp. 100–104). Briarcliff Manor, NY:

Academy of Management.

- McGivern, C. (1989). The dynamics of management succession: A model of chief executive succession in the small family firm. *Family Business Review*, 2(4):401-411.
- Neubauer, F. and Lank, A. G. (1998). *The Family Business*, Hampshire, UK: Palgrave Macmillan.
- Neubauer, F., & Lank, A. G. (1998). The family business: Its governance for sustainability. New York: Routledge.
- Pieper, T. M. & Klein, S. B. (2007) The Bulleye: A Systems Approach to Modeling Family Firms, *Family Business Review*, **20**(4):301-319.
- Piraeus Traders Association 2009. Family businesses have greater resistance toward the impacts of the financial crises. 19/03/2009
- Shanker, M.C., & Astrachan, J.H. (1996). Myths and realities: Family businesses' contribution to the US Economy- A framework for assessing family business statistics. *Family Business Review*, 9(2):107-123.
- Shepherd, D., & Zahra, S. (2003). From conservatism to entrepreneurialism: The case of Swedish family firms. Unpublished paper. Boulder, CO: University of Colorado.
- Singer, J., & Donoho, C. (1992). Strategic management planning for the successful family business. *Journal of Business and Entrepreneurship*, **4**(3):39-51.
- Stafford, K.; Duncan, K. A.; Danes, S., & Winter, M. (1999). A research model of sustainable family businesses. *Family Business Review*, **12**(3):197–208
- Steier, L. (2003).Variants of agency contracts in familyfinanced ventures as a continuum of familial altruistic and market rationalities. *Journal of Business Venturing*, 18(5):597–618.
- Suaré, K.C. and Santana-Martin, D.J. (2004), 'Governance in Spanish family business', *International Journal of Entrepreneurial Behavior and Research*, 10(1/2):141–63.
- Tagiuri, R. & Davis, J. (1996). Bivalent attributes of the family firm. *Family Business Review*, **9**(2):199–208.
- Ward, J. L. (1988). Keeping the family business healthy: How to plan for continuing growth, profitability and family leadership. San Francisco, CA: Jossey-Bass.
- Ward, J. L., (2010). *Keeping the Family Business Healthy*. Hampshire, UK: Palgrave Macmillan.
- Ward, J.L. (1987). Keeping the family business healthy: How to plan for continuing growth, profitability, and family leadership. San Francisco: Jossey-Bass.
- Whiteside, M. F. & Herz-Brown, F. (1991). Drawbacks of a dual systems approach to family firms: Can we expand our thinking? *Family Business Review*, 4(4):383–395.
- Zahra, Shanker A. (2005). Entrepreneurial Risk Taking in Family Firms. *Family Business Review*, **18**(1):23-40.

The journey of business location factors through years: a literature review

Harissia Vlachou¹, Olga Iakovidou²

Department of Agricultural Economics, Aristotle University of Thessaloniki, Greece ¹harig@agro.auth.gr ²olg@agro.auth.gr

Abstract Entrepreneurship has been widely considered as the power that drives development and well-being in the regions it affects. In other words, it is increasingly recognized as the vehicle that is improving the quality of life of individuals, families and communities. As a result, local, regional and national institutions as well as researches in different disciplines have focused on finding and researching the factors leading to the installation of business on a certain location. Research studies that have been published from time to time deal with the business location factors either by approaching individually one factor (i.e. the labor cost), or by studying a specific business category (i.e. biotechnology firms) or by researching the location where it is developed (i.e. urban). The scope of this article is to present past and present research studies related to the business location factors and have been occasionally examined and, showcase, at the same time, how they have developed in time. Thus, the presentation and documentation of the existing literature review could be a valuable tool for policy makers in regional and national level as well as for active entrepreneurs and researchers engaged in relevant academic disciples. Policy makers could take advantage of the business location factors so as to attract investments in regions seeking development while active entrepreneurs could realize the important role played by the factors affecting the location of their business in a specific place and researchers could trace and fill in potential research voids on this subject.

Keywords *business location decisions, business location factors, economic factors, quality of life factors.*

1 Introduction

As entrepreneurship has currently been an important element of economic development and innovation policy (Lafuente, Vaillant and Serarols, 2010), public and private economic development organizations, for many years, focused their efforts in a country or at a local level on attracting or nurturing new businesses, expanding existing businesses and retaining companies that are in situ. Where firms locate and why, has been a major question for researchers (Von Thünen, 1826; Marshall, 1890; Weber, 1929; Christaller, 1933; Hayter, 1997; Trullén, 2001; Parker, 2004; Autant-Bernard et al. 2006; Van Praag and Versloot, 2007; Ferreira et al., 2010;Lafuente et al., 2010) in economic geography, strategic planning, regional economics and organizational behavior for over fifty years (Kahn & Henderson, 1992).

Location theory has a long and rich historical tradition; its early development is associated, among others, with authors such as Weber (1929), Hoover (1948), Lösch (1954). Von Thünen (1826) was the first to conduct a study on this subject in view of the monthly rents of real estate ownerships. He claimed that distance is the most important factor in defining the price of the rent. Marshall (1890) focused on the agglomeration economies in view of industrial districts. Weber (1909) examined the factors influencing industrial location assuming that industrialists chose a location in view of cost minimization. Hoover (1948) outlined the spatial market-sharing agreement by combining the scale and agglomeration economies with the transportation cost.

After the unfolding of location theories, in the early 50s a growing body of researchers through empirical study boosts the investigation of the driving forces behind firms' location decisions. Yet, since then until nowadays much has changed, reclassification of markets, preferences of consumers, new environmental regulations, technological progress, evolution of transportation and renewed structural policies. In the same way, following all these changes the factors leading to the companies' location have evolved too.

The literature on business location factors can be broadly classified into categories: (a) studies measuring the influence of a specific factor or set of factors on firm location decisions; (b) studies explicating the location decision process for a specific industry or a business with specific characteristics and (c) studies identifying the location factors leading businesses in specific areas.

Examples of the first type include analyses of the impact of taxes, subsidies, and incentives (Buss 2001; Gius and Frese 2002; Hanson and Rohlin, 2011), environmental regulations (Bartik 1988; Brunnermeier and Levinson 2004), quality of life and amenities (Gottlieb 1995; Johnson and Rasker, 1995; Love and Crompton 1999; Granger and Blomquist 1999;Dissart and Deller 2000; Salvesen and Renski, 2003; Dahl and Sorenson, 2007; Kilvits, 2012), transportation and access (Krugman, 2001; Forkenbrock and Foster 1996; McQuaid et al., 1996; Bruinsma et al., 1997; Bryan et al., 1997; Button et al, 2005; Holl 2004a; Holl 2004b; Leitham et al., 2000; Targa, Clifton, and Mahmassani 2006) among others.

Research of the second type includes studies of the location decisions of biotechnology firms (Goetz and Morgan 1995; Feldman 2003; Koo, Bae, and Kim 2009; Su and Hung 2009), companies in the automobile sector (Bilbao-Ubillos and Camino-Beldarrain 2008; Klier and Rubenstein 2010), call centers (Bishop, Gripaios, and Bristow 2003; Richardson and Gillespie 2003) high-tech firms (Jarboe, 1986; Galbraith and De Noble, 1988; Frenkel 2001; Hackler 2003a,b, 2004) and many others. Even more specific are the studies of the location decisions of family firms (Kahn and Henderson, 1992); small enterprises (Sullivan et al., 1998; Liang et al., 2001).

Last but not least is the group of studies that connect location factors with special areas, like urban center (Cohen, 2000; Shukla and Waddel, 1991;) or rural areas (Johnson and Rasker, 1995; Yu and Artz, 2009; Lafuente et al., 2010; Vaillant et al., 2012).

As Kimelberg and Williams (2013) argue all types of research have clear implications for policy and practice. Data that illustrate how individual economic, social, or political factors affect the likelihood that firms will locate in a given place can inform policy decisions and economic development initiatives at the local, regional, state, and national levels. Similarly, a deep understanding of the array of factors that need to be present before a specific firm or industry will establish operations in an area is necessary to help government officials determine whether their municipality or region is—or could be—a viable candidate for such investment.

The purpose of this article is the timeless search and presentation of the literature on the factors business setup through empirical studies. More specifically, comparing the factors every 20 years as they are presented through the bibliography and distinguishing the most important ones are tasks of vital importance, for both the local authorities and government policies, as they can benefit from utilizing the findings of the researches in attracting new businesses and preventing the existing ones from leaving to other locations.

2 Employed Methodology

The core idea with a literature review is to summarize the state of art in the subject field, as a basis for identifying areas in which further research would be beneficial (Rowley and Slack, 2004). They state that literature reviews are important in: (i) supporting the identification of a research topic, question or hypothesis; (ii) identifying the literature to which the research will make a contribution, and contextualizing the research within that literature; (iii) building an understanding of theoretical concepts and terminology; (iv) facilitating the building of a bibliography or list of the sources that have been consulted;(v) suggesting research methods that might be useful; and (vi) analyzing and interpreting results. In conducting this literature review, we follow the general guidelines from Rowley and Slack(2004): (i) material collection, including (i) scanning documents, (ii) making notes, (iii)

structuring the literature review, (iv) building the bibliography, and (v) writing the literature review.

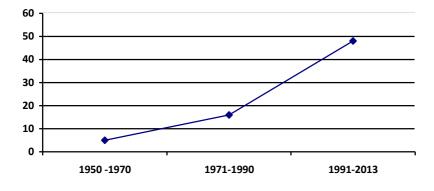
A comprehensive search of related articles from 1950 to 2013 was applied to produce a synthesis of peer-reviewed literature. The choice of the date 1950 as the starting point of the literature review was chosen for two main reasons: firstly, 50 years is a good timeframe to study the evolution of the location factors. Secondly, most search databases of scientific articles can be found from that period onwards. These 50 years, were divided into three groups of almost 20 years per each. It seems that this time is sufficient for the evolution of a phenomenon. The search strategy is based on selected databases (Scopus, Ebsco and Google scholar), and selected titles ("business" in combination with "location factors", "location decisions", or a combination of the following: "quality of life factors", "rural areas", "small firms". For example, one such combination was "quality of life factors + business location + rural areas". Also we used the terms "business/firms/plant location factors or business/firms/plant location decisions. Based on this and in the selected period 69 papers were mostly identified.

These 69 articles we reviewed are distinguished among 39 different international scientific journals. Seven (7) journals account for 26 articles (table I), while the other thirty one (31) articles are from 31 different journals. We also included three articles that are cited in two books and 9 papers that has been presented in Conferences. The articles of both cases, display highly in databases and are mentioned frequently in most of the articles. The highest number of articles is found in Economic Development Quaterly, Land Economics, Urban Studies, Growth and Change, Journal of Regional Science, Regional studies, Regional Sciences and Urban Economics and Technovation (Table 1).

Journal	Number of papers	Percentage (%)
Economic Development Quaterly	5	3.45
Land Economics	4	2.76
Urban Studies	4	2.76
Growth and Change	3	2.07
Journal of Regional Science	3	2.07
Regional Studies	3	2.07
Regional Sciences and Urban Economics	2	1.38
Technovation	2	1.38
Other	31	21.3
Chapters in Books	2	1.38
Papers presented in Conference	9	6.21

Table 1 Distribution of articles in Journals the period 1950-2013.

The literature of empirical surveys on location factors begins in the early 50s and until 1970 exhibit a low number of articles and a few years with no articles. The average number of publications during these first 20 years is 0.25 per year during this period. At this point, it must be noted that this particular period the scientists that has been occupied with this particular subject, preferred to write a book based in a theoretical perspective than presenting an article with empirical survey in a scientific journal. From 1971 until 1990 the number of articles starts to rise and the average number of publications is 0.8 per year. In the last period from 1991until 2013 there has been a baby boom in the study of business location factors and thee average number of papers is 2.08 per year.



3 Business location factors: determinants of site selection

Given the researches and theories that have appeared from time to time, an important number of researchers has put into groups the business location factors (Johnson and Rasker, 1995; Gottlieb, 1994; Van Dijk $\kappa\alpha$ Pellenbarg, 2000; Love and Crompton, 1999; Vailant, Lafuente and Serarols, 2012; Fernandes, Ferreira and Marques, 2012; Kimelberg and Nicoll, 2012; Kimelberg and Williams, 2013), while others (Baumbeck et al, 1973; Blair and Premus, 1987; Moore, Tyler & Elliott, 1991; Guimarães, Rolfe, and Woodward, 1998; Fisher $\kappa\alpha$ Peters, 1998) have focused either on the study of specific factors or of specific industry.

A classification of location factors has been made by Lloyd and Dicken (1977) followed by Van Dijk and Pellenbarg (2000) which distinguish them between : "firm internal" factors (e.g., quality of management, organizational goals, ownership structure, growth rate of turnover, employment and profits), "location" factors (absolute and relative characteristics of the location site, e.g. lot size and size of possible expansion space; distance to customers and suppliers), and "firm external" factors (e.g., government policy, regional economic structure, technological progress, etc.).

A totally different classification has been developed from Johnson and Rasker (1995), who examined the location decisions of firms to an amenity rich area of the Western USA. The four groups of location factors were: "Economic", "Qualitative", "Community and Recreation". More than four groups of location factors identify Karakaya and Canel (1998), they actually synthesize six: "cost" (with two component start up cost and cost of running a business), "quality of life" or "standard of living", "accessibility", "resources", "business envinroment" and "availability of existing building".

One year later, Love and Crompton (1999) put the variables related to the business location in five categories: "Quality of life", "Labor and cost issues", "Government involvement and taxes", "Daily living concerns" and "Proximity to relevant publics".

Many years after Lloyd and Dicken (1977) followed by Van Dijk and Pellenbarg (2000) classification, a similar approach has been chosen by Wardner (2012) categorizing the factors in three main groups: "Area factors" (community environment, market and competitors, cost of inputs, housing and recreation, physical characteristics, transportation and access). "Internal factors" with two sub-themes (firm: clients, employees, owners, profitability, supplier and work area to suit their specific requirement / building: physical conditions, image, facilities and amenities, flexibility, lease cost and tenure) and "external factors" (taxes and government incentives, costs and availability of utilities such as telecommunication cables and infrastructure, governance on a state and local level, environmental issues and the availability of both public and private capital).

More recently, two studies on Knowledge Intensive business location choices catalogued differently the factors. The first by Vailant, Lafuente and Serarols (2012) separate them in three groups: "infrastructure and economic motivations", "personal motivations" and "location-related motivations", while the second by Fernandes, Ferreira and Marques (2012) preferred four categories: "Economic conditions and local infrastructure", "Access to technologically superior knowledge", "Individual motivation" and "local characteristics".

Also, Kimelberg and Nicoll (2012), in their research on business location factors in the Medical device industry, grouped the 39 different factors into six broad topical categories: "labor", "permitting processes", "development and operating costs", "business environment", "transportation and access" and "quality of life/social environment". One year after the first author Kimelberg with Williams (2013) this time, grouped in the same way the locations in another research study.

3.1 The Period 1950 – 1970

As it can be argued the literature review on the business factors, has not been particularly rich these first twenty years. Articles in scientific journals are very measured while writing books about this theme richer. The interest of researchers focuses mainly on industry in different areas of USA and the factors that will lead to the choice of location.

At the beginning of this period Katona and Morgan (1952) tried through empirical research to clarify the factors influencing manufacturer's firm location in Michigan. They concluded that the major location factors were found to be: Personal reasons, to be near markets, availability of plants or sites. Ten years later in a review of various empirical studies Muller and Morgan (1962) found that traditional factors such as market access, labor costs, and raw materials were the most commonly mentioned by manufacturers.

Also in 1965 McMillan concluded in the following important factors with some minor variations. Markets tend to rank first or second. If the industries surveyed are resource oriented, raw materials will rank first and markets second. Market oriented industries will tend to rank markets first, labor second, and raw materials third. Transportation whether it is reported in terms of "central location to market" or "transportation facilities" will customarily rank third or fourth. One year later, in another empirical study of office decentralization, Wabe (1966) found the most important factors for the decentralization. These were expansion, integration of several offices, cost reduction and lease expiring

At the end of the first twenty years Logan (1970) also argues by far the most important locational consideration cited by manufacturers was access to markets, both industrial and consumer. Labor availability and skills were listed also as being particularly important. Industrialists indicated that they were deliberately seeking out rural areas of relatively low income (low wages) where there was a labor surplus: that is, the counties experiencing out-migration or high unemployment rates. The two other economic factors - the availability of land and buildings for industry and access to linked firms - would tend to attract firms to established industrial areas and possibly urban areas or communities from isolated rural districts. Also, a large proportion of the respondents listed "home area - personal reasons" as being important. Logan interprets this to mean industrialists locate in a particular area because they live there. Someone would therefore expect this factor to operate to the advantage of urban areas and industrial areas where most people likely to establish industries already live.

YEAR	AUTHOR/S	Study in Business Location (citation)	Market	Cost Factors	Labor	Raw materials	Transportation	Financial Incentives	Taxes	Government Regulations	Utuntes Personal Motives	Quality of life factors	Business Climate
1952	Katona & Morgan	Quantitative Factors (81)	✓								4		~
1962	Mueller & Morgan	Manufacturers (4)	√		√	✓							✓
1965	McMillan	Manufacturers (65)	√			√	√						
1966	Wabe	Office Decentralization (6)		1									~
1970	Logan	Industrial Plants in Winconin (1)	~		~	~					1		

Table 2: Surveys on Business Location Factors the period 1950-1970

3.2 The Period 1970 – 1990

During this period literature review begins to change regarding both the factors considered and the type of businesses. Industries are no longer the dominant topic of interest, as technology evolves and other forms of entrepreneurship thrive. At the same time, the usual industry location factors such as market access and labor give way to others, such as environmental regulations, trade unionism and quality of life. The issue of business location factors is just beginning to be of interest to both researchers and politicians alike.

3.2.1 Studying the "factor"

In the late '70s Foster (1977) was the first who studied the *economic and quality of life* factors in USA and Canada. A list of *economic factors* was presented and the one which was considered most important by respondents in USA from the survey was "dependability/productivity of workers". A second factor receiving a substantial percentage of respondents' choice was "wage rates". It should be obvious that productivity and wage rates are inter-related. In terms of labor, "availability of skilled workers" was also very important to respondents, and to a lesser extent "availability of semi-skilled/unskilled workers". Although not as important as the foregoing factors, "labor relations" were considered simply important. Factors as "nearness to customers" and "nearness to raw materials" were considered important by the Canadian businessmen.

In the same survey the respondents were requested to choose from *quality of life factors* those that are important to them personally, important to them in the daily operation of their plant and in terms of their perception of importance to their key personnel. The results were quite interesting as "good medical services" was the factor which rated highest in the individual personally category than in the other two. Also, "clean environment" was considered important in all three types of responses, and most particularly of the individual personally. "Good schools" are considered important for the respondent personally and for key personnel, but less important in terms of daily plant operation. Last the "streets safe from crime" seems to be one factor which figured prominently in all categories and most especially for the respondent personally.

Another factor was studied extensively by Fox (1981). In his research he argues that although there has been doubt about the importance of fiscal variables as determinants of industrial location, he has found that when areas which zone-out industrial firms are excluded from the sample, tax variables are statistically significant determinants of overall industrial location patterns. Also Wasylenko (1982) and Charney (1983) examined the role of *fiscal factors* in the intrametropolitan location decisions of manufacturing firms. Both of them emphasized on analyzing location choices of relocating firms, rather than existing patterns of economic activity or estimating changes in net employment. They resulted both that the property tax rate is

196

a significant location factor to relocating firms selecting sites within an urban area and that income tax differential has a negative but weak influence on those decisions. The location disincentives of property taxes are strongest for large firms and tend to decrease with firm size.

In 1985 Batrik searched the importance of unionization in business location decisions. The strongest conclusion of his is that differences in *unionization* across states are having a major impact on industrial location in the United States. The results also contradict the common view of those who argue that state and local *taxes* and public services exert no influence on business location patterns. This could have important implications for the incidence of state and local policies; as business location patterns change in response to taxes and services, land rents, local wages, and other prices will shift. In other words, he suggested that US states with lower wages and lower tax rates had higher odds of being chosen as sites for new plants in a study of the expansions of existing organizations.

Bartik (1988) also a few years later studied the relation between *envriromental regulations* and business locations. He concluded that there isn't any statistically significant effect of state environmental regulations on the location of new branch plants. The point estimates suggest that even sizable increases in the stringency of state environmental regulation are unlikely to have a large effect on the location decisions of the average industry. In the same conclusions, that the regulations didn't matter, resulted also the survey of McConnell and Schwab (1990), who studied the impact of environmental regulations specific to the motor vehicle industry location decisions.

3.2.2 Studying "The Business"

Dorf and Emerson (1978), have chosen to study the main determinants of *manufacturing business* in nonmetropolitan location. Findings occurred that the main factors were community size, distance from urban areas, and labor force. These determinants of plant location or expansion are independent of public action or control. Of secondary importance were property tax, railroads, and housing, of which two can be affected through public policy.

On the other hand in Epping's (1982) study the three most important factors to the *manufacturers*- locating-in-Arkansas group were labor, taxes, and industrial site. The three least important general factors were personal preferences, business services, and markets. The low ranking of markets was unexpected in light of earlier studies which found this factor to be very important. It could be that this factor is becoming less important over time.

Also Hekman (1982) studied the business location factors in North Carolina, South Carolina, and Virginia for the five-year period prior to 1982. His sample of 204 firms was largely composed of branch/plant operations of *multiplant firms* headquartered in the Northeast and Midwest. The business executives in the mail/telephone interviews were asked to rate the importance of 19 business location factors and 12 quality-of-

life factors on a scale of one to five. The significance of this approach is that it allowed for a comparison of the relative importance of business and quality-of-life location factors, at least for branch plant operations in the Southeast. The most important business location factors were state and local industrial climate, labor productivity, transportation, land availability (and room for expansion), and cost of land and construction. It is doubtful that land cost or availability would have induced the firm to locate in another region. Also, interestingly, markets did not appear as a significant factor in branch plant location choices. Last, it has been found that, the firms were more interested in a good business climate and low production costs (labor productivity). Quality-of-life factors were important in Hekman's survey study but their overall rating was less than the overall rating of traditional economic factors. The educational system, cost of living, housing, quality of air and water, and personal taxes were the top-ranked quality-of-life factors. Other quality-of-life factors such as climate, recreation, cultural resources, and entertainment received comparatively low ratings. Hekman's study also showed a rough degree of uniformity among different types of industries.

Hart, Denison and Henderson (1989) investigated directly the locational preferences of firms across different **industrial** sectors and various levels of technological sophistication. As the results of the majority of the same surveys the conclusions ended to be "proximity to markets/customers", "availability of space and facilities", "overall quality of the locality", "proximity to owner's or employee's residence".

Also, the same year, Lopez and Henderson (1989) studied the location decisions of 56 single-establishment, **small manufacturing food** processing firms in Mid-Atlantic states that were involved in processing vegetables, fruits, eggs, poultry, and/or seafood. Plant location decisions were found to be similar to the decisions of other manufacturing industries, i.e., input supply, product markets infrastructure, labor, and environmental regulations influence site selection. However, firms in their study were generally restricted to locations within commuting distance of the owner's residence.

Another type of businesses has been studied by Galbraith (1985). He concluded that *high technology firms* operate on a different set of factors from traditional industries in making their location decision. He identified five salient components which are the keys to high technology location decisions: the availability of professional and technical personnel; the general ambiance and lifestyle of the area; and the desire of the owner/CEO to live in the area, the climate and the community attitude towards business.

In an another effort to discover the location determinant of *high-technology firms* in the Ann Arbor area of Michigan, Jarboe (1986), addressed to forty six firms of small, rapidly growing new companies with a large percentage of their personnel devoted to research and development activities. The results of the survey concluded that access to area universities and the general quality of life, transportation networks and ability to attract and retain professional and skilled workers are the most important location factors for the high-technology firms in the area.

Similarly, a year later in 1987 a survey was conducted by Galbraith and De Noble (1988) in 226 *high technology* firms regarding location decisions. Results of the

survey suggested that higher technology firms are "footloose" in geographical location decisions, emphasizing the importance of ambiance and availability of labor and property. Results also suggested smaller firms place more emphasis on ambience, while larger ones emphasize business-related factors. For site-specific decisions, firms (smaller companies in particular) were influenced by cost and access factors. These findings are relevant for municipalities and other ancillary services interested in high technology development.

The above *state of empirical knowledge* regarding factors that influence location of firms have been sought by Blair and Premus (1987) in the end of the two decades. After reviewing the theoretical foundation of locational determinants, they described the decision process and discussed empirical findings with emphasis on findings from survey and econometric studies. Their conclusion from the review reveal that prior to the 1970's the conventional view of the dominant location factors were: access to markets, labor, raw materials and transportation. In more recent studies it is indicated that the traditional factors are still most important, but their dominance has been reduced as productivity, education, taxes, and community attitudes toward business and other factors have been recognized as influential. In other words the list of important locational determinants has been expanded with others noneconomic factors.

YEAR	AUTHOR/S	Study in Business Location (citation)	Market	Cost Factors	Labor	Raw materials	Transportation - Access	Financial	Incentives	Taxes	Government	Incgulations	Infrastructure	Personal Motives	Quality of life factors	Business Climate
1977	Foster	Economic & QoL Factors (21)		~	~	1									~	
1978	Dorf & Emerson	Manufacture in NonMetropolita n communities (45)	√		√		1			√						
1980	Wasylenko	Fiscal Factors (73)								√						
1981	Fox	Fiscal Factors (112)								✓						
1982	Charney	Fiscal Factors (92)								✓						
1982	Epping	Manufacturers (19)			√					√						√
1982	Hekman	Envinromental Regulation in Industry (Book)		~	√		1									~
1985	Batrik	Unionization – Taxes (634)								√						
1985	Galbraith	High-tech firms (5)			✓									√	~	√
1986	Jarboe	High-tech firms (17)			✓		~								~	
1987	Blair & Premus	Literature review (243)	~		✓	√	√			✓			√		√	~
<i>1988</i>	Batrik	Envinromental Regulations (219)									~					
<i>1988</i>	Galbraith & Noble	High –Tech Firms (62/book)		~	✓		~									~
1989	Hart, Denison & Henderson		√											√	✓	✓
1989	Lopez & Henderson	Small food manufucturers (30)	√		√						~		√			
1990	McConnell & Schwab	Environment al regulations in motor vehicle (186)									~					

 Table 3 Surveys on Business Location Factors the period 1971-1990

3.3 The period 1991 – 2013

The last two decades show great interest in the issue of business location factor. The list of articles in scientific journals abounds in various fields, related to the factors or the companies' divisions. To these areas of interest the site factor was added, with various studies focusing on urban or rural areas.

3.3.1 Studying the "Factor"

Transportation

The role of transport has a long tradition in classical location theory. Several survey studies (Krugman, 1990; Button et al., 1995; Forkenbroch & Foster, 1996) examined the relation between *transportation* and location decision of firms. For example new economic geography models emphasize the importance of transport costs along with imperfect competition, market size and economies of scale in explaining the location of industry (Krugman 1991). In 1995, Button et al., estimated the importance of transport (road links, bus links, air links and rail links) included and related infrastructure in the location/relocation process. The analysis showed that road links were considered by a firm's representatives as the most important factor in company location. The survey concluded that weak transportation infrastructure does not stimulate the firm's migration, but the nature and quality of the transportation system become a very important factor for companies' decisions to locate/relocate.

Among empirical location studies, several have adopted a survey approach for the relation between **transportation** and business location decisions (McQuaid et al 1996; Bruinsma et al 1997; Bryan et al 1997; Leitham et al 2000). Such studies generally involve a series of questions relating to firm characteristics and various location factors including issues of transport. Among the various types of transport infrastructure, roads are frequently reported as the most important type.

Although firms perceive the availability of good **transport** infrastructure as very important, it is seldom the decisive factor in a location decision. The findings of such survey research show that new transport infrastructure seems to relax location constraints, allowing firms to examine a wider range of locations. Nevertheless, most location decisions and relocations are made over very short distances, and firms are often found to relocate within the same region (Bruinsma et al 1997).

The range of potential benefits to firms due to **transport** improvements has been highlighted by Bryan et al. (1997) in their survey of local firms following the construction of the A55 motorway in North Wales. Their findings suggest important logistical impacts. Most firms are reported to have benefited in terms of reduced input ordering times, improved output delivery and customer service. Whilst the findings reported above are useful, the main drawback of the survey approach is that the results are restricted to the particular research context and cannot easily be generalised, nor do they provide explanations of why a particular behaviour of a firm has been observed, i.e. a particular ranking of location factors.

Also, Holl (2004a) studied the impact of *road infrastructure* investment on the location of new manufacturing establishments in Spanish municipalities from 1980 to 1994 using Poisson regression. These results showed that new motorways affect the spatial distribution of manufacturing establishments by increasing the attractiveness of municipalities close to the new infrastructure. Holl (2004b) reports similar results for manufacturing plant location in Portugal.

Similar to the previous studies, Arauzo (2005) also found that in Catalonia the location of industry is effected by *transportation and accessibility*, especially highways which are important determinants for the location of firms and more so than trains.

Apart from previous studies several others have recognized the extent to which *transportation investments* may improve access to markets and, by doing so, influence firms' location decisions. Such an empirical study developed by Targa et al. (2006) in the geographical area of the region in Maryland. The analysis established significant association between transportation supply and firm-level relocation decision in the study area and underscores the role of other firm and area- of-influence attributes in this process. The findings suggest a positive association between access to primary highway facilities and the level of economic activity. Also the results confirm expectations that roads with higher functional form and capacity are likely to be more influential in the location and relocation decisions of businesses.

Another factor has been investigated by Fotopoulos and Spence (1999), who proxy the availability of *infrastructure* by the amount of public investment per capita as an explanatory variable in a regression model accounting for new plant openings in Greece. They find a significant and positive effect of public infrastructure on new plant openings. These findings are supported by Holtz-Eakin and Lovely (1996) who test the impact of *public infrastructure* in a general equilibrium context and find that impacts of public infrastructure primarily work through increases in the number of manufacturing establishments.

In 2004 the researchers Brunnermeier and Levinson offer a review and critique of the large literature on the pollution havens hypothesis: "the state of environmental regulations in the early literature, based on cross-sectional analyses, typically conclude that they have an insignificant effect on firm location decisions. However, more recent studies that use panel data to control for unobserved heterogeneity, or instruments to control for endogeneity, find statistically significant pollution haven effects of reasonable magnitude. Furthermore, this distinction appears regardless of whether the studies look across countries, states, counties, or industries, or whether they examine plant locations, investment, or international trade patterns".

Financial Incentives- Taxes

Atikis et al (2011) in an attempt to determine the degree of relevance of industrial location policy to the location criteria used by manufacturing firms, identified 16 location factors which are of importance in Greece, regardless of geographical area, type of business or the respondents' functional responsibility. The main purpose of the research was to assess *financial incentives* as a motivation for business attraction

in particular areas. The result showed that government financial incentives are not of particular importance in the selection of the location for a plant in Greece, so there is a need to reevaluate its approach towards regional economic development.

Almost ten years later Polyzos and Minetos (2007) in their study in regional development incentives and their influence on industrial location decisions report on their findings: "The results on direct and indirect population potential suggest that interregional transportation infrastructure is of outmost importance for industrial location. Other transportation infrastructures such as ports and airports may not be of the same importance as road infrastructure. On the other hand, the provision of suitable locations for industrial uses is an important location factor. In addition, proximity to metropolitan areas or to large urban concentration seems to be highly influential on the behaviour of firms. Incentives set by the state are important location factors but not always and everywhere. In certain localities, other factors might be more important. The tradition of a region in accommodating on activities of the secondary sector is also an important location force. Hierarchy and human capital need to be interpreted very carefully. As regards hierarchy it is often ranked relatively low by companies when seeking suitable location. The complex indicator of human capital may need further investigation and may also require distinguishing particular manufacturing activities within the secondary sector".

In the beginning of 2000, Buss (2001) collected and reviewed the *tax* study literature to assess the state of knowledge about the relationship among taxes, related factors, and economic growth as well as the use of tax incentives to influence business locations. In the end he concluded by arguing that tax literature, now in hundreds of publications, provides little guidance to policy makers trying to fine-tune economic development. Taxes should matter to states, but researchers cannot say how, when, and where with much certainty. Firms may need tax incentives to increase their viability in some locations, but researchers cannot definitively say which businesses or which locations.

Next year the study of Gius and Frese (2002) attempted to determine the effect of state personal and corporate *tax rates* on firm location. Using a random effect model on a data set of 14,000 observations, the present study finds that a state's personal tax rate has a negative effect on firm location but that a state's corporate tax rate has no statistically-significant effect on firm location. These results suggest that the locational decision of a firm is affected more by the impact that high personal taxes will have on the manager's and/or owner's income than by the negative impact of high corporate tax rates on firm profits.

A wide range of state and local governments use tax incentives as an economic (re)development tool. Part of the hope of policy makers is to attract new establishments to the local economy. Hanson and Rohlin (2011) examined how offering **tax incentives** in a local area affects the entry of new business establishments. They resulted that there was a positive and statistically significant effect of a specific tax incentive program on attracting new establishments, a result that is particularly strong in the retail and service industries.

Quality of Life – Personal motives

In 1994 Gottlieb (1994) tried to capture *quality of life* as a determinant factor in a firm's location decision. He actually reviewed the theoretical, survey and econometric literature on amenity oriented firm location and employment growth. The results were quite suprising, as envinromental quality for both categories (all firms and high-Tech firms) ranked first as the main factor for business location. As he moved from all category firms to high-tech category school quality, cultural amenities and public safety dropped in importance. On the contrary, cost of living, cost of housing and community issues rose in importance for high-tech firms.

A year later, Johnson and Rasker (1995) studied the *economic and quality of life* values and suggested that in the familiar view of business location values, such as tax structure and cost of doing business, it should be added also other values that may be important to the business location decisions, such as the role of a quality environment, scenic beauty, low crime rate, and recreational opportunities. In their empirical study, the findings -over 500 firms in Greater Yellowstone Region between long-time resident business (old-timers) and relative newcomer business owners (newcomers)-indicate that values that reflect the quality of the living environment are important to the business location decision by rural business owners. More specifically, those who have lived in the study region for more than five years tended to rate the importance of the quality of life values more highly than did newcomer business owners.

Love and Crompton (1999) tried to capture from key decisions –makers the role of *quality of life* elements in the location of 174 businesses that had relocated, expanded and been launched in Colorado. Five domains of elements were derived and the set of quality of life elements ranked in importance behind labor and cost issues and daily living concerns. As the authors argue, quality of life was most important to companies that moved into Colorado outside the state, were relatively footloose, had fewer than eight employees and employed a high proportion of professionals.

Another study the same year by Granger and Blomquist (1999) investigates the notion of amenities in influencing manufacturers' location choices in urban areas. The authors suggest that if amenities affect wages, land values and other costs, then amenities will influence location decisions. Using urban, county-level, Census data and regression models they estimated the location of small and medium-sized manufacturing establishments. Holding constant scale and agglomeration economies, amenities, measured by a quality-of-life index, are found to influence manufacturers' location with the effects varying by industry. Labor-intensive industries are more strongly attracted to high-amenity urban locations.

In the beginning of the decade Dissart and Deller (2000) presented an overview of the literature on the topic of Quality of Life. Among the topics reviewed was the quality of life as a determinant factor for business location decisions. In their conclusions they mention characteristically: "Subsidy incentives are not particularly significant predictors of firm location. And, as the United States shifts to a more

service-based economy, the overall importance of nontraditional locational factors will increase in significance while traditional locational factors decline in significance. This is not to say that the latter do not lay a role: traditional factors (land, labor, capital, infrastructure, location) are vital ingredients for economic development. It is only after these basic factors are satisfied that one may turn to more intangible factors such as quality of life to increase the competitive edge of a place".

"The importance of *Quality of life* in the location decisions of new economy firms" is the exact title of Salvesen and Renski's (2003) research. They have conducted a pilot study to examine the actual location decisions of a small sample of firms that have recently located in Raleigh, Durham and Chapel Hill metropolitan area of North Carolina. The researchers argue that none of the respondents cited quality of life as being the most important factor in their business location decision. Several firms stated that quality of life was one of the important factors including cost of land and the quality and cost of labor.

Regarding personal motives, Dahl and Sorenson (2007) argue that *social capital* places strong constraints on an entrepreneur's ability to found a firm in a region in which he or she does not have connections. The main purpose of their thesis was to investigate whether entrepreneurs tend to open and locate businesses in regions in which they have deep roots ('home' regions). Also, they further investigate if their ventures perform better (survive longer) when they locate in these home regions. The results ended with the outcome that the value of social capital moreover appears substantial, similar in magnitude to the value of having prior experience in the industry entered (i.e. human capital).

Finally, according to Kilvits (2012) nowadays soft factors such as "quality of life" (housing, environment and infrastructure), "image" of places or "private" reasons are important determinants of firm locations. The climate, low crime, educational system, cost of living, quality and cost of housing, quality of air and water, recreation facilities, etc. (all modern living and work environment) are very important for potential high-technology investors and skilled labor.

3.3.2 Studying the "Business"

During these twenty years, the studies have focused on the examination of the establishment factors of individual firms and have been about businesses of a specific entrepreneurial activity, of a certain size and of a particular administrative organization.

In the late '90s Crone (1997) identified variables related to *manufacturing* firms' business locations decisions, without examining the influence of firm size. The variables examined in his study were accessibility to the market, facilities, the region cost of some important inputs, labor costs, energy prices, and taxes.

Another attempt remarkable was a study of 87 Australian *SMEs manufacturing* in Adelaide undertaken by Kupke and Pearce (1998). In this survey the authors found that the two most important industrial location factors for owner-managers were being

close to the central business district (CBD) and having direct access to main roads. This finding appears to be similar to those of large firms. On the contrary the approach of *small manufacturing* enterprises in Vermont considering the key factors influencing location decisions was made by the application of a nonparametric analytical procedure. The findings of the study seem to indicate that small manufacturers' location decisions are often related to personal factors including environment (quality of life) and local residence (want to stay home). Also important business factors were related to finance (access to capital) markets (customers in the local and regional area) and the availability of facilities (Liang, et al, 2001).

The objective of Henderson and McNamara (2000) study was to identify local characteristics influencing the location of new *food manufacturing* plant investments. In the study, the county characteristics have been analyzed and associated with the location of food processing plant investments. The results of the survey concluded in these major factors: access to input and product markets, agglomeration economies, access to a transportation system, low wages, and local tax policies are factors that influence food manufacturing investment locations.

Much of the research has focused on the decision-making processes of large firms; however some attention has been given to how SMEs make such decisions. For example, an attempt has been made from Moore, Tyler and Elliot (1991) who identifies the key factors in the small and medium enterprises (*SME*) location decision from a survey of almost 1400 companies. The most important factor was the availability of regional development assistance, followed by the quality and size of the labor supply, including wage levels, and the potential for future expansion. Infrastructure was found to be relatively unimportant as a locational determinant. These factors apply with considerable uniformity to different industrial sectors, but there are major differences between countries within the European Community (EC).

Also Sullivan, Halbrendt and Buescher (1998) studied the *small business* location considerations and extendedly conclude that there are different factors affecting location decisions, between firm size. So large-size firms in comparison with small and medium size firms, place most importance on physical infrastructure, such as access to the interstate, situated on a freight bearing highway, access to railroads, access to a commercial airport, port or harbor facilities. Also, large firms compared to the others two categories, place significantly greater importance on the availability of managerial/professional workers, availability of unskilled labor, availability of mass transportation for workers, favorable local labor costs, availability of low cost commercial loans and availability of development assistant. As for quality of life factors, small firms considered low relative to their location decision significantly more than other size firms. On the other hand, large firms considered availability of social services (such as hospitals) relative to their location decision significantly more than other size firms.

Another approach has been developed by Coughlin and Segev (2000). They examined the county-level pattern of new *foreign-owned manufacturing* plants in the United States from 1989 through 1994. They concluded that economic size, educational attainment, the existing manufacturing base, and transportation infrastructure are found to be positive, statistically significant determinants of

location. In addition, foreign investors tend to prefer urban locations. Meanwhile, higher average labor-intensiveness and higher taxes as a share of gross state product are found to deter foreign direct investment. A surprising finding is that foreign investors tend to locate in counties with higher percentages of black population.

As it is known in all decades industrial location attracts the interest of the scientific community regardless of the country and continent the researcher comes from. One such study is the Badri's (2007) effort to approach the dimensions of **industrial** factors. In his research he identified fourteen critical factors utilizing judgments by previous scientific authors and a group of industrial location professionals. These fourteen critical factors are transportation, labor, raw materials, markets industrial sites, utilities, government attitude, tax structure, climate and community. But Badri argues that five detailed factors did not meet the criteria set for being "critical" or "important". These factors were availability of postal services, worker stability, adequacy of sewage facilities, availability of religious facilities, and availability of library facilities.

Jensen & Pompelli (2002) examined the perceived importance of site location characteristics identified in a 1999 survey of 198 small Tennessee agribusinesses. As they characteristically express "Responding firms ranked proximity to buyers/customers, labor, and raw materials above other factors. However, the relative importance of all factors varied by industry subsector. For example, compared to food processing firms, textile milling and lumber/wood products firms perceived community incentives as less important. Projected firm growth and current location also affected the perceived importance of site location factors. The diversity of perceived factor importance across agribusiness subsectors supports the idea that incentives and promotion of site location factors to attract small agribusiness may need to be tailored to meet specific firms' needs".

Kahn and Henderson (1992) study the location preferences of *family and non family firms* and they conclude that both categories ranked proximity to costumers and markets first, suggesting that business success is the primary concern regardless of ownership form. It is also interesting than while family firms are more concerned with proximity to residence, there are not with other quality-of-life items. On the other hand, nonfamily firms prefer locations that minimize facilities and employment cost while providing access to skilled labor and access to public and private research facilities.

Recently the location decisions of *Knowledge Intensive Business Services (KIBS)* and Knowledge *Intensive Service Acitivies (KISAs)* became a major topic for many researches mainly in Spain and Portugal. Example of such a case is the empirical study of 500 KIBS in Portugal (Fernandes, Ferreira and Marques, 2012). From a set of twenty nine factors explaining the choice locations of such firms, three have been found to be of most significance: economic conditions and local infrastructures, access to technologically superior knowledge, individual motivations and local characteristics.

A more specified research developed Kimelberg and Nicoll (2012) by studying 48 *Medical* device *firms* in Massachussetts. The findings of the survey clearly emphasize that the availability of appropriate labor is the single most important factor driving site selection for medical technology firms. Also the availability of on-site parking, the timeliness of approvals and appeals, the crime rate in the local area and the state tax/financial incentives round out the top five most highly rated factors.

3.3.3 Studying the "Area"

Urban Areas (Cities)

Against previous studies which deal with business site selection decisions mainly concerned in cost factors, Karakaya and Canel (1998) attempt to provide empirical evidence on the importance of cost but also other location related variables among 84 fastest – growing businesses in *New England and New York*. Availability of skilled labor, transportation facilities, state tax rate, state regulatory environment and real estate tax were the major factors that the survey concluded in influencing site selection.

As Cohen (2000) argues the business sector is a fundamental determinant of a firm's location choice. The rankings vary when businesses are classified by industry sector, corporate function, size or technological intensity. Retail and personal service businesses locate to maximize sales revenue rather than to minimize transportation costs. The location of retail and personal service firms are largely dictated by existing and anticipated patterns of target residential populations, particularly affluent households with greater disposable income. There is also a tendency for specialty retail and personal service establishments to cluster in specialty shopping or entertainment districts, typically in or near the downtowns of major urban centers. Such agglomerations of similar activities tend to attract more patrons then if they were spread across a larger area.

In 2006 Prat and Marcén reported an examination of factors that have influenced Spanish companies in Aragon in their choice of business locations. In the province of the city Zaragoza 129 industries were surveyed to serve as the primary source of information. The results have been interesting as the high number of customers in the area ranked first with small difference with personal/subjective factors, such as factory's proximity to the entrepreneur's residence and the business being original to the area were given the next highest importance. Other factors which also mattered in varying degrees were: proximity to supply sources in general, industrial density of the area has industrial tradition in the sector, proximity to densely populated areas, availability of qualified industrial soil, the existence of good financial centers, the existence of skilled labor and urban infrastructure.

Rural Areas

Firms created by locals (rural local entrepreneurs) and their location decision, was the main subject of the Michelacci and Silva (2005) study. As they support, family background and social factors affect entrepreneurs' location decision making. The experience of growing up in a particular community with a certain entrepreneurial climate or culture can impact both an individual's decision to become an entrepreneur and their decision about business location. Rural residency experience may shape attitudes toward rural entrepreneurship and provide rural entrepreneurial role models.

As Yu and Artz (2009) argue it is commonly known business located in *rural areas* will suffer from less local demand, infrastructure, or supportive assistance than in urban areas, but will benefit from lower land rents and wages. Starting a business in a rural or urban area is an important but complex decision to make. Multiple dimensions deserve attention in rural entrepreneurship and rural development: business nature, human capital and expertise, family background and social capital. In their empirical study they conclude that entrepreneurs from origins tend to choose to start businesses in rural areas. Social capital and social networks established in one's home region are shown to be a strong factor in the location choice of entrepreneurs. Half of the entrepreneurs migrate back to their home state after graduation, likely to take advantage of local networks or due to familiarity with local comparative advantages.

Serarols, Vaillant and Urbano (2009) analyzed the establishment of high and medium technology-based manufacturing firms (HMTBMFs) specifically in *rural areas*. The survey was conducted in 34firms and resulted contrary to the most literature, that the location decision of technology-based entrepreneurship in rural areas is not a fruit of calculative and rational economic thinking, nor is it apparently swayed by the potential benefits that may come from institutional spillovers or from public incentives. The conclusions coming from the sampled entrepreneurs point toward a decision-making process that is more emotional, revolving around the entrepreneurs' desire to establish residence or remain part of the rural community where they located their businesses. More specifically the interviewed rural Catalan HMTBMF's search for a specific lifestyle and quality of life dominated in their business location decision-making process.

In another study the variables that influence the choice of location made by *rural and urban* knowledge intensive activity firms (KISA), have been the major purpose of Lafuente, Vaillant and Serarols (2010) study and therefore they extensively studied them. The findings of that survey indicated that in the case of the surveyed new Catalan KISAs, the search of the lifestyle and quality of life dominated the business location decision-process. In other words, the personal motivation factors, such as the entrepreneur's personal motives and the search of quality of life attracted KISA's towards rural areas. This on the contrary was not found to be the case for urban-based new created KISAs, as their entrepreneurs placed more importance on aspects related with local attitudes to entrepreneurship and their business, or to the availability of physical infrastructure. Two years later, in 2012 the same authors Vaillant, Lafuente et Serarols (2012) presented in their article the same results.

YEAR	AUTHOR/S	Study in Business Location (citation)	Market	Cost Factors	Labor	Raw materials	Transportation – Access	Financial Incentives	Taxes	Government Regulations	Infrastructure	Personal Motives	Quality of life factors	Business Climate
1990	Krugman	Transportation (8162)					~							
1991	Moore, Tyler & Elliot	Regional Incentives & Infrastructure			√			~						√
1992	Cahn & Henderson	Family and Nonfamily firms (23)	√		√							√	√	
1994	Gotlieb	Transportation Access (12)											√	
1995	Johnson & Rasker	Economic& QoL (146)											√	
1995	Button et al	Transport (49)					√							
1996	Forkenbrock & Foster	Highways (29)					✓							
1996	McQuaid, Eitham & Nelson	Accessibility (25)					✓							
YEAR	AUTHOR/S	Study in Business Location (citation)	Market	Cost Factors	Labor	Raw materials	Transportation – Access	Financial Incentives	Taxes	Government Regulations	Infrastructure	Personal Motives	Quality of life factors	Business Climate
1996	Holtz-Eakin & Lovely	Public Infrastructure (110)									✓			
1997	Bruinsma, Rienstra, & Rietveld	Highways (50)					√							
1997	Bryan et al	infrastructure (12)					√							
1997	Crone	Manufactures (15)	✓	✓					✓					
1998	Karakaya &Canel	Businesses in Cities (49)			√		1		√	1				

Table 3 Surveys during the period 1991-2013

1998	Suillivan, Halbredt & Buescher	Small Businessess			✓		1						~	
1999	Granger & Blomquist	Amenities- Quality of Life (31)	,										√	
1999	Love & Crompton	Quality of Life (46)		√	√							✓	✓	
1999	Fotopoulos & Spence	Infrastructure (60)									√			
2000	Leitham, McQuaid & Nelson	Transport (62)					√							
2000	Cohen	Cities (49)			✓		~				✓		✓	
2000	Coughlin & Segev	Foreign-owned manufacturing (223)					~							√
2000	Dissart & Deller	Quality of Life (120)											√	
2000	Kupke & Pearce	SME manifactures	1				~							
2000	Henderson & McNamara	Food Manufactures (27)	✓		√		√		✓					
		Literature												
2001	Buss	Review of Tax incentives (165)							✓					
2001 YEAR	Buss AUTHOR/S	Review of Tax incentives	Market	Cost Factors	Labor	Raw materials	Transportation – Access	Financial Incentives	Taxes	Government Regulations	Infrastructure	Personal Motives	Quality of life factors	Business Climate
		Review of Tax incentives (165) Study in Business Location	 ▲ Market 	✓ Cost Factors	Labor	Raw materials	Transportation – Access	Financial Incentives	Taxes	Government Regulations	Infrastructure	✓ Personal Motives	 ≺ Quality of life factors 	Business Climate
YEAR	AUTHOR/S	Review of Tax incentives (165) Study in Business Location (citation) Small Manufacturing Enterprises in	 ✓ Market 	✓ Cost Factors	Labor	Raw materials	Transportation – Access	Financial Incentives	 ✓ Taxes 	Government Regulations	Infrastructure	✓ Personal Motives	 ≺ Quality of life factors 	Business Climate
YEAR 2001	AUTHOR/S Liang et al	Review of Tax incentives (165) Study in Business Location (citation) Small Manufacturing Enterprises in Vermont Tax incentives	 ✓ Market 	✓ Cost Factors	Labor	★ Raw materials	Transportation – Access	Financial Incentives		Government Regulations	Infrastructure	★ Personal Motives	 ≺ Quality of life factors 	Business Climate

2004	Brunnemeier & Levinson	(158)							√					
2004a	Holl	Manufacturing firms in Portugal (61)					~							
2004b	Holl	Manufacturing firms in Spain (147)					✓							
2005	Arauzo	Industry in Catalanian Municipaliies (95)					~							
2005	Michelacci & O'Silva	Local Entrepreneurs (124)										√		
2006	Prat & Marcén	Business in Zaragoza (1)	√		~						~	√		√
2006	Targa & Clifton	Transportation Access (12)					1							
2007	Polyzos & Minetos	Regional Incentives in Greece	√				1	1						
2007														
2007	Badri	Industry	√	√	✓		√		✓	√	✓			✓
2007	Badri Dahl & Sorenson	Industry Social Capital (29)	~	~	~		√		~	~	√	✓		✓
	Dahl & Sorenson	Social Capital	Market	Cost Factors	Labor	Raw materials	Transportation –	Financial Incentives	Taxes	Government Regulations	Infrastructure	Personal Motives	Quality of life factors	Business Climate
2007	Dahl & Sorenson	Social Capital (29) Study in Business Location	Market	Cost Factors	Labor	Raw materials	Transportation –	Financial Incentives	Taxes	Government Regulations	Infrastructure	 ✓ Personal Motives 	 ≺ Quality of life factors 	Business Climate
2007 YEAR	Dahl & Sorenson AUTHOR/S Yu & Artz	Social Capital (29) Study in Business Location (citation) Rural Entrepreneurship		Cost Factors 🔸	Labor	Raw materials	Transportation –	Financial Incentives	Taxes	Government Kegulations	Infrastructure	 ✓ ✓ ✓ Personal Motives ✓ 	 ✓ ✓ Quality of life factors 	Business Climate

2011	Attikis, Merikas & Vozikis	Financial Incentives in Greece (1)		√	
2011	Hanson & Rohlin	Tax incentives (8)		\checkmark	
2012	Kimelberg & Nicoll	Medical Device Industry (1)	~	√ √	~
2012	Vaillant, Lafuente & Serarols	Rural Knowledge Intensive Service Activities (KISA)			\checkmark \checkmark
2012	Fernandes, Ferrera & Marques	Knowledge Intensive Business Services (KIBS)			√ √ √
2012	Kilvits	Living Envinroment			1
2013	Kimelberg & Williams	Influence of Facility type	√	√ √	1

4 Conclusions

The study of the business location factors in an area, according to the literature review, has been the topic of research of many different scientists and science alike for over 50 years. Its importance appears to be increasing over the years and to be varying as it evolves.

Therefore, ever since 1950 until today, researchers from different continents have either been studying the effect that specific factors have on business establishment in a place or have been looking for business location factors in the establishment of specific type of companies. Of course, in many studies both cases are combined during research, e.g. the role of quality of life in high-technology firms location decisions.

Over the first twenty years (1950-1970) the empirical studies on the business location factors, presented in scientific journals' articles, are few and derive mainly from research in various areas of America. The industry and the factors that shape its decision to establish are the dominant entrepreneurial activity being studied. As

expected, the survey results in three key variables: proximity to markets, labor and raw materials.

Another point worth mentioning about this period is that the term used most often in database articles is the word 'plant' instead of 'business' or "firms'. One could characterize this era as "industrial location economically oriented".

The next twenty years (1971-1990) witnessed a growth of empirical studies dealing with "business location factors", particularly in the second decade. These empirical studies are divided into two categories: a) those researching the effect of a particular factor on a business location and b) those researching a particular category of businesses.

In the latter category, the interest of researchers in manufactures seems to prevail, while at the same time others appear to be really interested in high-technology firms. Moreover, it is apparent that while certain factors, like distance from markets, labor and raw materials have an impact on manufactories, other minor factors are added in this period to the first ones, such as state and local taxes, community attitudes toward business, and transportation. This is further reinforced by the study of Jungthirapanich and Benjamin (1995) which conducted a hierarchy of eight location factors in industry location starting with "market" as the most important, followed by labour, site consideration, raw materials, transportation, and services utilities, government concerns and in the end "community environment" as the least important factor.

Unlike manufactories, high-technology firms are influenced by different factors such as: availability of professional and technical personnel, the general ambience and lifestyle of an area as well as the desire of the owner to live in an area.

Regarding *the first category*, studying a specific business location factor, the ones that stand out are: fiscal factors, unionization, environmental regulations and quality of life factors, with the latter one to be gaining more and more ground during the past twenty years.

The term "plant" used in most articles appears to be decreasing while the term "business" begins to appear dynamically. This era could be characterized as "a transition from economic factors to noneconomic factors in business location".

Over the last 22 years, since 1991 onwards, the study of business location factors has been of immense interest to many researchers both in America and Europe. Many a researcher has been studying individual specific factors (transportation, taxes, etc) to investigate the influence of each one separately on business location.

Thus, in the early 90s there is a large number of articles focusing on the study of transportation and accessibility factors, sometimes combined and at other times synthetically. Most scholars conclude that the existence of transport infrastructure (motorways, ferries and airports), or the means of transport and accessibility are essential factors in many cases. Worth quoting is the article by Krugman (1991) that has 8,162 citations in Google Scholars' database, showing of what great importance the transportation factor is for researchers and maybe politicians.

Other recently studies relating to the importance of "public infrastructure" or "environmental regulations" in business location decisions, reach the same conclusions with "transportation" factor. Similar results derive from empirical studies examining the factor "financial incentives", which seems not to be a powerful one either.

As far as "tax" factors are concerned, it is not clear whether it has a major effect on a business establishment; therefore, conclusions vary according to each case separately. The truth is that high taxation of individual and business income has a negative impact on attracting business investment.

Last but not least, the "Quality of life" factor appears to hold all the more particular interest for researchers who investigate which variables (environment, climate, crime safety, schools, cultural amenities, etc.) are likely to lead businesses to a site location. The truth is that for some business categories (e.g. high-technology firms) this factor is the most influential one so for the businessman to settle in a place as for the set up of his business.

Regarding the study of business, industries during this period continue to appear in articles with unflagging interest until the late 90s. This time, the influence of location factors appears in different sizes and industry sectors, suggesting that the classical factors are less significant, while the importance of non-economic factors is gradually rising.

The same conclusions are drawn in articles on numerous studies in recent years on Knowledge Intensive Service Activities (KISA) and Knowledge Intensive Business Services (KISA).

What is worth mentioning about this period is that more recently there has been a shift of interest from researchers from the study of factors and that of businesses to the study the site (urban or mostly rural areas) and its features to the business location decisions. The findings indicate that "quality of life" factor has played a major role for both cities and particularly rural areas.

In these last twenty-two years the term "plant" doesn't appear very often in the scientific articles, while the terms "business" or "firm" are most used by authors. This era could be characterized as "The increasing importance of quality of life factors".

To sum up, the main conclusions drawn from this literature review are the following:

- 1. The study of business location factors was and still remains the subject of study of many different disciplines. Moreover, as Schmenner, Huber, and Cook, (1987) claim '...determining factors which influence a company's location has remained an elusive quest' and to date Paleti, Bhat and Singh (2012) '...the choice of a location to start a new business or to expand into new locations for an existing business is critical to the success of the entity making such decisions'.
- 2. None of the researchers end up in their studies in the same conclusions, even when they investigate the same industries, the same business categories or the same factors. Blair and Premus (1987) emphasized that no two studies reveal identical findings because of regional differences, evolving conditions of production, industry-specific versus more general studies, and new firms versus expansion of existing ones.

- 3. The latter leads to the conclusion that the decisions taken by entrepreneurs vary depending on the type of business, the place and the criteria chosen as to where they will settle.
- 4. The literature seems to be driven from economic factors (market, labor and raw materials) to non-economic factors (transportation, quality of life).
- 5. Finally, the results of various research projects reveal that no conclusions can be drawn or generalized. In this way, the entrepreneurs' decisions leading to a business location are not always the same and are not to be taken for granted._

References

- Arauzo, J. M. (2005). Determinants of industrial location: An application for Catalan municipalities*. Papers in Regional Science, 84(1), 105-120.
- Artikis, G. P., Merikas, A., & Vozikis, G. S. (2011). Revisiting The Use Of Financial Incentives As An Instrument Of Regional Economic Development Policy: The Case Of Greece. Journal of Applied Business Research (JABR), 12(4), 89-99.
- Autant-Bernard, C., Mangematin, V., Massard, N. 2006. Creation of Biotech SMEs in France.
- Badri, M. A. (2007). Dimensions of Industrial Location Factors: Review and Exploration. Journal Of Business And Public Affairs, 1(2),
- Bartik, T. J. (1985). Business location decisions in the United States: Estimates of the effects of unionization, taxes, and other characteristics of states. Journal of Business & Economic Statistics, 3(1), 14-22.
- Bartik, T. J. (1988). The effects of environmental regulation on business location in the United States. Growth and Change, 19(3), 22-44.
- Baumback, C., Lawyer, K. and Kelly P. (1973). "How to organize and operate a Small Business", (5th Edition), Englewood Clifts, NJ: Prentice-Hall, Inc. pp. 137-176.
- Bilbao-Ubillos, J., and Camino-Beldarrain V. (2008). "Proximity matters? European Union enlargement and relocation of activities: The case of the Spanish automotive industry". Economic Development Quarterly, 22(2),149–166.
- Bishop, P., Gripaios P., and Bristow G. (2003). "Determinants of call centre location: Some evidence for UK urban areas", Urban Studies, Vol. 40, No 13, pp. 2751–2768.
- Blair, J. P., & Premus, R. (1987). Major factors in industrial location: A review. Economic Development Quarterly, 1(1), 72-85.
- Bruinsma, F. R., Rienstra, S. A., & Rietveld, P. (1997). Economic Impacts of the Construction of a Transport Corridor: A Multi-level and Multiapproach Case Study for the Construction of the A1 Highway in the Netherlands. Regional Studies, 31(4), 391-402.
- Brunnermeier, S. B., & Levinson, A. (2004). Examining the evidence on environmental regulations and industry location. The Journal of Environment & Development, 13(1), 6-41.

- Bryan, J., Hill, S., Munday, M., & Roberts, A. (1997). Road infrastructure and economic development in the periphery: the case of A55 improvements in North Wales. Journal of Transport Geography, 5(4), 227-237.
- Burdina M., (2004). "Impact of transportation on business location decisions in rural upper great plains". MSc thesis, Major Department Agribusiness and Applied Economics, North Dacota State University of Agriculture and Applied Science.
- Buss, T. F. (2001). The effect of state tax incentives on economic growth and firm location decisions: An overview of the literature. Economic Development Quarterly, 15(1), 90-105.
- Button, K. J., Leitham, S., McQuaid, R. W., & Nelson, J. D. (1995). Transport and industrial and commercial location. The Annals of Regional Science, 29(2), 189-206.
- Charney, A. H. (1983). Intraurban manufacturing location decisions and local tax differentials. Journal of Urban Economics, 14(2), 184-205.
- Christaller, W., 1933. Die Zentralen Orte in Südeuschland (in Deutch). Wissenschaftlische Buchgesellschaft, Darmstad.
- Cohen, N. (2000). Business location decision-making and the cities: bringing companies back. Brookings Institution Center on Urban and Metropolitan Policy.
- Coughlin, C. C., & Segev, E. (2000). Location determinants of new foreign-owned manufacturing plants. Journal of regional Science, 40(2), 323-351.
- Crone, T. M. (1997). Where Have All the Factory Jobs Gone—and Why?. Business Review, 136(1), 3-18.
- Dahl, M. S., & Sorenson, O. (2007, October). Home sweet home: Social capital and location choice. In Proceedings of 2007 DRUID Summer Conference.
- Deller, S., Lledo V., and Marcouiller D. (2008). "Modeling regional economic growth with a focus on amenities", Review of Urban and Regional Development Studies Vol. 20, No 1, pp. 1–21.
- Dijk, J. Van, & Pellenberg, P. H. (1999). Firms spatial dynamics of firm behavior (p. 262)
- Dissart, J. C., & Deller, S. C. (2000). Quality of life in the planning literature. Journal of Planning Literature, 15(1), 135-161.
- Dorf, R. J., & Emerson, M. J. (1978). Determinants of manufacturing plant location for nonmetropolitan communities in the west north central region of the US. Journal of Regional Science, 18(1), 109-120.
- EPPING, G. M. (1982). Important factors in plant location in 1980. Growth and Change, 13(2), 47-51.
- Feldman, M. (2003). "The locational dynamics of the US biotech industry: Knowledge externalities and the anchor hypothesis", Industry and Innovation, Vol. 10, No 3, pp. 311–328.
- Fernandes, C., Ferreira, J., & Marques, C. (2012). Knowledge Intensive Business Services Entrepreneurship and factors driving entrepreneurs's decisions in choosing firm location. In Proceedings of USASBE, page 1018.
- Ferreira, J., Marques, C., Fernandes, C. 2010. Decision-making for location of New Knowledge Intensive Businesses on ICT sector: Portuguese Evidences, International Journal of E-Entrepreneurship and Innovation 1(1), 60-82.

- Fisher, Peter S., and Peters Alan H. (1998). "Industrial Incentives: Competition among American States and Cities", W. E. Upjohn Institute: Kalamazoo, Michigan.
- Forkenbrock, D. J., & Foster, N. S. (1996). Highways and business location decisions. Economic Development Quarterly, 10(3), 239-248.
- Foster, R. (1977). Economic and quality of life factors in industrial location decisions. Social Indicators Research, 4(1), 247-265.
- Fotopoulos, G., & Spence, N. (1999). Spatial variations in new manufacturing plant openings: some empirical evidence from Greece. Regional Studies, 33(3), 219-229.
- Fox, W. F. (1981). Fiscal differentials and industrial location: Some empirical evidence. Urban Studies, 18(1), 105-111.
- Frenkel, A. (2001). "Why high-technology firms choose to locate in or near metropolitan areas", Urban Studies, Vol. 38, No 7, pp. 1083–1101.
- Galbraith, C. S. (1985). High-Technology Location and Development: The Case of Orange County. California Management Review, 28(1), 98-109.
- Galbraith, C., & De Noble, A. F. (1988). Location Decisions by High Technology Firms: A Comparison of Firm Size, Industry Type and Institutional Form. Entrepreneurship: Theory & Practice, 13(2), 31-47.
- Gius, M. P., & Frese, P. (2002). The impact of state personal and corporate tax rates on firm location. Applied Economics Letters, 9(1), 47-49.
- Goetz, S., and Morgan R. (1995). "State-level locational determinants of biotechnology firms", Economic Development Quarterly, Vol. 9, No 2, pp. 174–184.
- Gottlieb, P. (1994). Amenities as an economic development tool: Is there enough evidence?. Economic Development Quarterly, 8(3), 270-285.
- Gottlieb, P. (1995). "Residential amenities, firm location and economic development". Urban Studies, Vol. 32, No 9, pp. 1413–1436.
- Granger, M. D., & Blomquist, G. C. (1999). Evaluating the influence of amenities on the location of manufacturing establishments in urban areas. Urban Studies, 36(11), 1859-1873.
- Green, G. (2001). "Amenities and community economic development: Strategies for sustainability", Journal of Regional Analysis and Policy, Vol. 31, No2, pp. 61–75.
- Guimarães, P., Figueiredo, O. and Woodward D. (1998). "Regional Incentives and industrial location in Puerto Rico", International Regional Science Review, Vol. 21, No 2, pp. 119-138.
- Hackler, D. (2003a). "High-tech growth and telecommunications infrastructure in cities". Urban Affairs Review, Vol. 39, No 1, pp. 59–86.
- Hackler, D. (2003b). "High-tech location in five metropolitan areas", Journal of Urban Affairs Vol. 25, No 5, pp. 625–640.
- Hackler, D. (2004). "The information technology industry and telecommunications: An empirical analysis of cities in the Minneapolis-St. Paul and Phoenix metropolitan areas". Journal of Urban Technology, Vol. 11, No 3, pp. 35–59.
- Hanson, Andrew, and Shawn Rohlin. "DO LOCATION-BASED TAX INCENTIVES ATTRACT NEW BUSINESS ESTABLISHMENTS?*." Journal of Regional Science 51.3 (2011): 427-449.

- Hart, S. L., Denison, D. R., & Henderson, D. A. (1989). A contingency approach to firm location: the influence of industrial sector and level of technology. Policy Studies Journal, 17(3), 599-623.
- Hayter, R., 1997. The dynamics of industrial location: the factory, the firm and the production system. New York: Wiley
- Hekman, J. S. (1982) "Survey of Location Decisions in the South." Economic Review, pp. 6-19.
- Henderson, J. R. & McNamara, K. T., 2000. "The Location Of Food Manufacturing Plant Investments In Corn Belt Counties," Journal of Agricultural and Resource Economics, Western Agricultural Economics Association, vol. 25(02), December.
- Holl, A. (2004a). Manufacturing location and impacts of road transport infrastructure: empirical evidence from Spain. Regional Science and Urban Economics, 34(3), 341-363.
- Holl, A. (2004b). Start-ups and relocations: Manufacturing plant location in Portugal*. Papers in Regional Science, 83(4), 649-668.
- Holtz-Eakin, D., & Lovely, M. E. (1996). Scale economies, returns to variety, and the productivity of public infrastructure. Regional Science and Urban Economics, 26(2), 105-123.
- Hoover, E., 1948. The location of economic activity . McGraw Hill, New York.
- Jarboe, K. P. (1986). Location decisions of high-technology firms: a case study. Technovation, 4(2), 117-129.
- Jensen, K., & Pompelli, G. (2002). Manufacturing site location preferences of small agribusiness firms. Journal of Small Business Management, 40(3), 204-218.
- Johnson, J. D., & Rasker, R. (1995). The role of economic and quality of life values in rural business location. Journal of Rural Studies, 11(4), 405-416.
- Jungthirapanich C. and C.O. Benjamin, (1995). "A knowledge-based decision support system for locating in manufacturing facility", IIE Transactions, Vol. 27, No. 6, pp. 789-99.
- Kahn, J. A., & Henderson, D. A. (1992). Location Preferences of Family Firms: Strategic Decision Making or "Home Sweet Home"?. Family Business Review, 5(3), 271-282.
- Karakaya, F., & Canel, C. (1998). Underlying dimensions of business location decisions. Industrial Management & Data Systems, 98(7), 321-329.
- Katona, G., & Morgan, J. N. (1952). The quantitative study of factors determining business decisions. The Quarterly Journal of Economics, 66(1), 67-90.
- Kilvits, K. (2012). Living Environment as Location Decision Factor for Manufacturing Enterprises. China-USA Business Review, 11(2), 217-224.
- Kimelberg, S. M., & Nicoll, L. A. (2012). Business Location Decisions in the Medical Device Industry Evidence From Massachusetts. Economic Development Quarterly, 26(1), 34-49.
- Kimelberg, S. M., & Williams, E. (2013). Evaluating the Importance of Business Location Factors: The Influence of Facility Type. Growth and Change, 44(1), 92-117.
- Klier, T., and J. Rubenstein (2010). "The changing geography of North American motor vehicle production", Cambridge Journal of Regions, Economy and Society Vol. 3, No 3, pp. 335–347.

- Koo, J., Bae, J., and Kim D. (2009). "What does it take to become a biotech hot spot?" Environment and Planning. C, Government and Policy, Vol. 27, No 4, pp. 665–683.
- Krugman, P. (1990). Increasing returns and economic geography (No. w3275). National Bureau of Economic Research.
- Kupke, V. and J. Pearce, (2000)."Identifying Industrial Location and Site Preferences for Small Business", Pacific Rim Property Research Journal vol. 6, no. 1, pp. 12-23 1324-1443
- Lafuente, E., Vaillant, Y., & Serarols, C. (2010). Location decisions of knowledgebased entrepreneurs: Why some Catalan KISAs choose to be rural?. Technovation, 30(11), 590-600.
- Leitham, S., McQuaid, R. W., & D Nelson, J. (2000). The influence of transport on industrial location choice: a stated preference experiment. Transportation Research Part A: Policy and Practice, 34(7), 515-535.
- Liang, Chyi-lyi (Kathleen), Yeong-tzay Su, Paul Dunn, and Catherine Halbrendt,(2001), A Nonparametric Approach to Study Key Factors Influencing Location Decisions for Small Manufacturing Enterprises in Vermont, presented at the second Annual USASBE/SBIDA Joint National Conference, February 7-10, 2001.
- Lloyd, P. & Dicken, P. (1977). Location in Space: a Theoretical Approach to Economic Geography. New York: Harper and Row. Dijk, J. Van,
- Logan, M. I. (1970). Locational decisions in industrial plants in Wisconsin. Land Economics, 46(3), 325-328.
- Lopez, R. A., & Henderson, N. R. (1989). The determinants of location choices for food processing plants. Agribusiness, 5(6), 619-632.
- Losch A (1954) The economics of location. Yale University Press, Connecticut
- Love, L. L., & Crompton, J. L. (1999). The role of quality of life in business (re) location decisions. Journal of Business Research, 44(3), 211-222.
- McConnell, V. D., & Schwab, R. M. (1990). The Impact of Environmental Regulation on Industry Location Decisions: The Motor Vehicle Industry. Land Economics, 66(1), 67.
- McMillan, T. E. (1965). Why manufacturers choose plant locations vs. determinants of plant locations. Land Economics, 41(3), 239-246.
- McQuaid, R. W., Leitham, S., & Nelson, J. D. (1996). Accessibility and location decisions in a peripheral region of Europe: A logit analysis. Regional Studies, 30(6), 579-588.
- Michelacci, C., & Silva, O. (2007). Why so many local entrepreneurs?. The Review of Economics and Statistics, 89(4), 615-633.
- Moore, B., Tyler, P., & Elliott, D. (1991). The influence of regional development incentives and infrastructure on the location of small and medium sized companies in Europe. Urban Studies, 28(6), 1001-1026.
- Mueller, E., & Morgan, J. N. (1962). Location decisions of manufacturers. The American Economic Review, 52(2), 204-217.
- Paleti, R., Bhat, C. R., & Singh, P. (2013). A Spatial Multivariate Count Model for Firm Location Decisions. In Transportation Research Board 92nd Annual Meeting (No. 13-4253).

- Parker, S.C., 2004. The Economics of Self-Employment and Entrepreneurship . Cambridge University Press.
- Polyzos, S. and D. Minetos (2007). "Regional Adevelopment incentives and their influence on the location decisions of industrial firms in Greece: An ordinal regression analysis", Discussion Paper Series, Vol. 13, Vol.8, pp. 193-224, Department of Planning and Regional Development, School of Engineering, University of Thessaly.
- Prat, J. C. R. S., & Marcén, R. F. (2006). Influential factors in location choice of Spanish businesses in Aragon. Journal of Entrepreneurship, 15(1), 63-81.
- Richardson, R., and Gillespie A. (2003). "The call of the wild: Call centers and economic development in rural areas", Growth and Change Vol. 34, No 1, pp. 87–108.
- Rowley, J., Slack, F., 2004. Conducting a literature review. Management Research News 27 (6), 31–39
- Salvesen, D. and Renski H. (2003). "The importance of quality of life in the location decisions of new economy firms", Center for Urban and Regional Studies, Economic Development Administration, U.S. Department of Commerce.
- Schmenner, R. W., Huber, J. C., & Cook, R. L. (1987). Geographic differences and the location of new manufacturing facilities. Journal of Urban Economics, 21(1), 83-104.
- Schmenner, Roger W. (1980). "Choosing New Industrial Capacity: On-Site Expansion, Branching, and Relocation," Quarterly Journal of Economics Vol. 95, No 1, pp. 103–119.
- Serarols, C., Vaillant, Y., & Urbano, D. (2009). Rural Technology-Based Entrepreneurs. In Public policies for fostering entrepreneurship (pp. 245-271). Springer US.
- Serarols, C., Vaillant, Y., Urbano, D., 2009. Why would a technology-based entrepreneur locate in a rural area? Answers from rural Catalonia. In: Leitao, J., Baptista, R. (Eds.), Public Policies for Fostering Entrepreneurship, (pp. 245-271), Springer, New York.
- Shukla, V., & Waddell, P. (1991). Firm location and land use in discrete urban space: a study of the spatial structure of Dallas-Fort Worth. Regional Science and Urban Economics, 21(2), 225-253.
- Small Business Economics 26, 173-187.
- Su, Y., and Hung L. (2009). "Spontaneous vs. policy-driven: The origin and evolution of the biotechnology cluster", Technological Forecasting and Social Change Vol. 76, No 5, pp. 608–619
- Sullivan, P., Halbrendt, C., Buescher, M. "Small Business Location Considerations for Agriculture and Forestry Sectors." Proceedings of the International Council for Small Businesses 43rd World Conference, Singapore, June 8 -10, 1998.
- Targa, F., Clifton, K. J., & Mahmassani, H. S. (2006). Influence of transportation access on individual firm location decisions. Transportation Research Record: Journal of the Transportation Research Board, 1977(1), 179-189.
- Trullén, J., 2001. Léconomia de Barcelona: cap a un nou model de desenvolupament. Revista Situación Catalunya , 2, 26-38.

- Vaillant, Y., Lafuente, E., & Serarols, C. (2012). Location decisions of new 'Knowledge Intensive Service Activity'firms: the rural–urban divide. The Service Industries Journal, 32(16), 2543-2563.
- Van Dijk, J. and Pellenbarg P. (2000). "Firm relocation Decisions in the Netherlands: an ordered logit approach", Regional Science, Vol. 79, No 2, pp. 191-219.
- Van Praag, M.C., Versloot, P.H., 2007. What is the value of entrepreneurship? A review of recent research. Small Business Economics 29, 351–382
- Wabe, J. S. (1966). Office Decentralisation: An Empirical Study. Urban Studies, 3(1), 35-55.
- Wardner, P. (2012). Understanding the role of "Sense of place" in office location decisions. 18th Annual Pacific-Rim Real Estate Society Conference, Adelaide, Australia, 15-18 January 2012.
- Wasylenko, M. J. (1980). Evidence of fiscal differentials and intrametropolitan firm relocation. Land Economics, 56(3), 339-349.
- Weber A (1929) Theory of the Location of Industries. University of Chicago Press, Chicago
- Yu, Li and G. Artz (2009), "Migrationn and Rural Entrepreneurship", Working Paper No 09017, Department of Economics, Iowa State University.

Information and Communication Technologies

Using PKI to increase the security of the electronic transactions

Marin Aranitasi¹, Marsida Ibro², Bojken Shehu³

¹Center for R&D, Polytechnic University of Tirana maranitasi@fti.edu.al
²Department of Electronic and Telecommunication, Polytechnic University of Tirana marsida_ibro@hotmail.com
³Department of Informatics Engineering, Polytechnic University of Tirana bojken.shehu@gmail.com

Abstract. One of the most critical elements of security is the ability to provide trust to the electronic transactions. One technology that helps injecting trust and confidence to these transactions is Public Key Infrastructure (PKI). Today, PKI has been viewed as critical and necessary not only to the commercial sector but also to the government sector.

The purpose of this paper is the demonstration of a clear and stable method to increase the security of electronic data transactions. This method was implemented in a government communication system. We had analyzed the actual problems in the electronic system of the Albanian Government and we propose a new and stable solution.

Keywords: PKI, Electronic Transaction, Security, Digital Certificate.

1 Introduction

The today world is totally dependent from e-commerce. Internet is the most common and effective tool to increase the internet transactions. Someone will say that we have the tools for affording trust to internet. But this is not totally true. For example antiviruses and firewalls do not help in establishing the identity of the parties during transactions. Also these solutions do not help in the understanding of the level of trust when we are doing an online transaction or dealing with an individual.

So we need a method or a technology to provide trust and confidence to the internet. Some of the few technologies that can accomplish this include **Public Key Infrastructure** (**PKI**). For a lot of people, PKI includes complex and long deployments. Perhaps this may have been real a few years ago, but today is different. As we mentioned before, PKI has been viewed as a critical and very necessary not only to the private sector, but also to the government sector. A lot of countries, within the last few years have passed laws that make digital signatures legally equivalent to physically drafted signatures. In addition, many countries also have some regulatory elements to the Certificate Authorities to ensure the quality of their operations. In the second section we will analyze the problems that have the Albanian government these days. In the third section we mention some of the national PKI government projects that we take into consideration. In section 4 we give our schematic solution and then we explain our detailed scheme. In section 5 are the conclusions and we present our future plans and in section 6.

2 **Problem description**

Government services are divided in three major groups:

- 1. Government to Citizens (G2C)
- 2. Government to Business (G2B)
- 3. Government to Government (G2G)

Actually in Albania the Government to Business group of services is operational. Each of these services has their respective web sites. This is a good thing, but there is a problem with the number of websites addresses to remember. Also the users have to remember the credential for accessing each web site. This is quite impossible. The problems can be summarized as follows:

- 1. Users have to remember a lot of identification elements
- 2. Users in case of a lost or in case they forget the id elements, have to go to the specific institution, with an official request, to get back their id information.
- 3. Every institution has to create help desk structures that in 80-90% face with issuance of id elements.
- 4. This mechanism of management has big problems, because we can't guarantee the authenticity of the operations with the electronics services, if the credentials are so "OPEN".

We can divide these problems in two main categories:

- 1. Citizen problems: waste of time, going from one institution to another, waiting, and probably not solving the problem in time.
- 2. Institution problems: have increased costs :
 - The institution that gives the information has paper and printing costs, because the electronic information must be printed.
 - The institution that accepts the information has equipment costs, because the information is in paper format and needs to be entered in the system.

The objective of this project is:

To design an electronic communication infrastructure between the electronic systems of the different institutions, that allows the online identification of their users.

3 Related works

There are many government initiatives that are based on PKI. According to [1] one of the biggest government PKI-based projects in the United States is "Common Access Card (CAC)", developed by the Department of Defense (DoD). CAC holds specific information and applications on PKI certificates, for the particular needs of the department. From the perspective of PKI, one of the CAC smart card holds up to three certificates to keep personnel information.

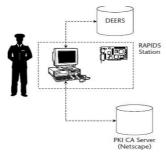


Fig. 1. CAC

Three certificates stored in the CAC include:

- 1. An authentication certificate. This certification is reserved for transactions. Specifically in this certificate is used to sign documents and to access secure DoD Web portals.
- 2. A signature certificate. This certificate is used to sign e-mail.
- 3. An encryption certificate. The certificate is used by others to send encrypted e-mail to users of CAC card.

In the following chart we see the dependencies between the authentication and price.



Fig. 2. Authentication vs Price

As seen from this graph identification technologies begin to be confident when starting to use "authentication with two elements".

Another project that we take into consideration is the South Korean national PKI project [2].

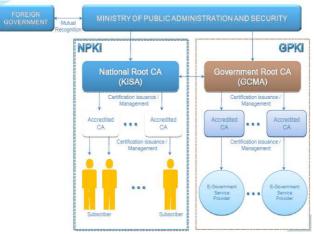
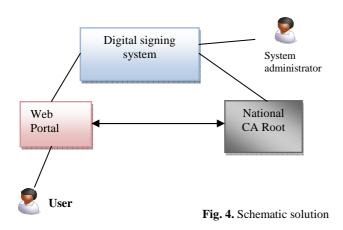


Fig. 3. South Korean National PKI project

These projects show scalability and reliability under extreme environmental conditions such as high security, with a very large population.

4 SCHEMATIC SOLUTION



Given that in Albania there are actually NO systems, that resembles the previously mentioned systems, namely CAC and Korean system, we are presenting a solution that is based the on basic principles of PKI, but this solution is adapted to the conditions and specifications of our country. The theoretical solution to the problems we mentioned in section 2, is presented in Figure 4.

The explanation of terms in the above chart is as follows:

- The user the service user
- **SCDev** the equipment that contain the citizen electronic identity and digital signature.
- Web portal. Portal in which the services are offered on-line.
- National Root CA. PKI infrastructure which issues certificates for electronic signature. This infrastructure also publishes the CRL list (list without valid certificates) which is used for OCSP service. Also this infrastructure will provide TSP service.
- OCSP. Online Certificate Status Protocol [3]
- **TSP** Time stamping.[4]
- CRL. Certificate Revocation List[5]
- **E-Signature System**. The application that signs the document (including the processes for determining the validity and time stamp).
- **The system administrator**. Person responsible for the administration of the system.

So we need to get into more detail to explain our solution

4.1 Hardware architecture of the solution

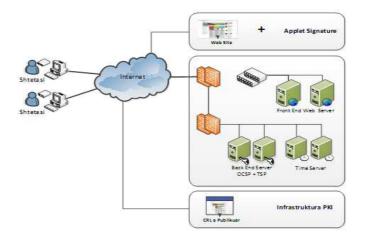
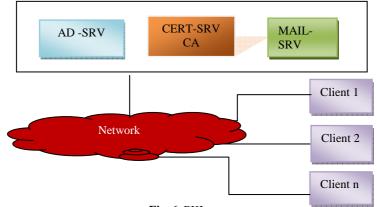


Fig. 5. Hardware architecture

Based on the fact that a small number of signatures will be realize in the first year of implementation, in order to have an acceptable economic solution, we decided that we will implement OCSP and TSP services on the same server. This solution is also provided with full redundancy in each of the components. Also the solution is conceived to be scalable in each of its components, in order to handle the increasing number of requests for service.

The components of the system are:

- Web Site. This component is the portal in which will be uploaded the documents that need to be signed electronically (e.g. procurement portal, tax portal, etc.). An application for signature must be installed in this portal. This application needs to integrate several of its components with the existing structure of the web-site
- **PKI infrastructure**. This is a PKI infrastructure of an existing or new CA that will be established. For our system is of primary importance to co operate with two components:
 - To have access to the published CRL list, which is used from OCSP system, to validate the certificate of the user.
 - To confirm the identity of the user (service requester).
- OCSP / TSP. This component of the system serves to implement the signature-based services (time stamp application and the application for validation and electronic signature). The system is designed with two levels (tier) of hardware. In the first level we have placed two Web Servers, after the firewalls that protect servers from the Internet and in load balance. In the second level we have placed 2 servers, to execute the OSCP and TSP applications. In each of the servers will be installed Hardware Security Module (HSM) and a local data base. We will install 2 time servers connected to a GPS antenna. These servers will provide the exact time that will be used for the TSP application.



4.2 Public Key Infrastructure

We decided to build practically one part of the proposed system. Our construction was carried out in a virtual environment with a limited number of users. The introduction of certificates in electronic communications increases the security of the system, especially when it comes to delicate as that government systems.

As given in Fig. 5, one of the three elements of the scheme is the PKI infrastructure. Fig. 6 shows what we had done. It consists of:

- Active directory Server (AD-SRV)
- Certificate Authority Server (CERT-SRV CA)
- Mail server (MAIL-SRV)
- Clients

Using [6][7][8][9][10] we set up our Certificate authority. We built also a mail server. We decide that our certification authority will issue two kinds of certificates.

- A login certificate
- A signing certificate

We connected the active directory server with Certification Authority. We also created some users. For each user we issue a digital certificate that contains his credentials. For security reasons we decide to set the time validity to one year. We put the certificate into a smart card. To have a secure but not expensive system we decide to use as device that will keep the certificates of the users, that is a smart card. This means that each user needs a smart card of its own to hold his/her digital certificate. Each smart card needs a smart card reader to be used. Now the login process in the system, for each user, is done only with the establishment of its smart card in a reader device and he can immediately use the system, without having to remember his credentials. Then we connected the CA with the mail server. That means that we "told" to the mail server that the outgoing emails must be signed by the CA. The certificate used for this work is the signing one. Tests went very well which means that emails that were sent by the mail server were signed by the certificate generated by the CA. This mechanism simplifies a lot the logging process and secures mail correspondence. We decide to issue only two types of certificates because the services that are now operational are focused only on digital signing and remote login in to the system.

5 Conclusions

We saw in this paper the problems that have actually the Albanian government in the process of digitalization of their processes and the identification of their users. To solve these problems we propose to build a PKI infrastructure. But this infrastructure was adapted to our standards. Today in Albania the group services that are available is the government to business. This group needs to send and receive secure documents with the government institutions. We presented our conceptual solution and then we present our detailed solution to increase the security of these transactions. The scheme of solution was divided into three little projects. The first was the set up of the PKI infrastructure. We set up this infrastructure. The core of it is the CA certification authority. Also we decide to issue two types of certificates the login certificate and the

signing certificate that electronically signs the documents and e-mails of the business users. So the problem of identification of the users is solved. We need now to fully construct our system, i.e. to built the OCSP and TSP servers and integrate all of them into the web site. Now the Albanian government is a "little" more secure.

6 Future work

We divide the hardware solution in three individual projects. The first part, the PKI infrastructure is done successfully. The second (the OSCP&TS) and the third, the web site, are going to be implemented in these months. As we know the lightweight directory access protocol (LDAP) is the Internet standard way of accessing directory services that conform to the X.500 data model.[11] When implementing these projects we will take into consideration the deficiencies when LDAP is used to support PKI that were presented in [12]. We suppose that we will use the new method proposed in [13][14]. Also we will take into consideration the problems between different PKI implementations and the solution given to these problems by [15]. In the end we are going to implement the web site that will be the interface between the users and the system. Details of these implementations we are going to present in another paper.

References

- 1. Wiley.:PKI Security Solutions for the Enterprise Solving HIPAA, E-Paper Act, and Other Compliance Issues, 160--163, (2003)
- 2. Jaejung Kim.: Overview of National PKI Establishment, April (2009).
- M. Myers, R. Ankney, A. Malpani, S. Galperin and C. Adams, *Online certificate status* protocol – OCSP, Internet Engineering Task Force: RFC 2560, June 1999.
- 4. Open TSA available at: http://opentsa.org/
- 5. R.Housley, W.Polk, W.Ford, D.Solo.: Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile RFC 3280, April (2002)
- 6. Open CA LibPKI, available at: https://www.openca.org/projects/libpki/
- 7. Open CA OCSPD, available at: https://www.openca.org/projects/ocspd/
- 8. Open CA project homepage, available at: https://www.openca.org/
- 9. Open SLP project, available at: http://www.openspl.org
- 10. Open SSL homepage, available at: http://www.openssl.org/
- 11. K. Zeilenga, Ed. Open LDAP foundation.:*Lightweight Directory Access Protocol* (*LDAP*): *Technical Specification Road Map*", RFC 4510 June (2006)
- 12. David Chadwick, University of Salford.: *Deficiencies in LDAP when used to support Public Key Infrastructures*" May (2001)
- Jong Hyuk Choi, Sang Seok Lim, and Kurt D. Zeilenga.: A New On-line Certificate Validation Method using LDAP Component Matching Technology. Workshop on Information Assurance and Security United States Military Academy, West Point, NY (2005)

- Sang Seok Lim, Jong Hyuk Cho,i Kurt D. Zeilenga.: Design and Implementation of LDAP Component Matching for Flexible and Secure Certificate Access in PKI. April (2005)
- 15. Massimiliano Pala and Sean W. Smith.: *Finding the PKI needles in the Internet haystack. Computer Science Department, Dartmouth College, Sudikoff, Hanover, NH, USA,* Journal of Computer Security, (2010)

Simulation of Network Request Mapping Techniques

Adelina Basholli¹, and Thomas Lagkas¹,

¹ Computer Science Department, The University of Sheffield Internationaly Faculty, CITY College, Thessaloniki, Greece { abasholli, tlagkas }@city.academic.gr

Abstract. Wireless broadband networks are designed to provide high quality services to concurrent mobile users, using the Orthogonally Frequency Division Multiple Access schema for frame structure. This schema defines rectangular resource allocation of time slots and frequency carriers. This structure is used for arranging the incoming user requests. Aiming to increase bandwidth utilization while arranging the incoming user requests into an effective way, we have considered and evaluated Bin-Packing algorithms. This paper presents analysis and design of various Bin packing algorithms, including a new developed algorithm, named Guillotine First Fit Algorithm and a new version of Shelf First Fit algorithms with different input values are gathered, analyzed and compared. The results are encouraging and provide indications regarding the usage of particular algorithms in practice.

Keywords: OFDMA, Bin Packing algorithms, Bandwidth utilization.

1 Introduction

Furious rate of technological change reflected in growth in mobile communication technologies. These transformations made very challenging and widely spread mobile communication functionalities, especially with the usage of mobile phones as communication tools [1]. This is evident while considering the generations of mobile communication and their improvements from generation to generation [2- 4]. Provided services, generally include: data transfer, international communication - roaming, multimedia services and voice over IP services, e-commerce, global positioning system services, and many other related services [5-8]. Therefore, mobile communication technologies are integral part of human's everyday life and they are developing day by day, by offering new services which categorizes them in four generations [9] (until now) of mobile communication technologies. This development

required different changes in both radio access part1 of the networks and in the core part2, too [10]. Concerning the radio access part, various multiplexing techniques have been developed, such as: Time Division Multiple Access (TDMA), Frequency Division Multiple Access (FDMA), Code Division Multiple Access (CDMA), Wide Code Division Multiple Access (WCDMA), or Orthogonally Frequency Division Multiple Access (OFDMA).

Regardless the generation of mobile communication technologies, the prior aim is the user satisfaction with the quality of offered services and the ability to be always ready to support different user requirements. In order to achieve these objectives, large capacity of bandwidth is required. Large bandwidth capacity implies higher data rates, support of more services and fulfillment of more user requirements in an efficient way. User requests present subscriber's requests for additional services from service providers (telephony services, internet services, etc). In order to accomplish these needs, various multiplexing techniques may help. One of them is OFDMA that is used in fourth generation (4G) technologies [11]. OFDMA is known as a technique that defines a rectangular resource space (time by frequency) in which the user requests are arranged.

A recurring problem that telecommunication industry confronts is the usage of a method which will increase bandwidth utilization while organizing incoming frames³ in an efficient way. Towards this problem, different methods have been suggested. Especially the usage of Non-deterministic Polynomial-time hard algorithms, respectively two dimensional packing algorithms, is recommended [12- 14]. This is briefly presented in section 2.

The purpose of this paper is twofold. Our first aim is to present a new version of Shelf First Fit Algorithm and a combined-new algorithm which we have named it Guillotine First Fit Algorithm. These algorithms are part of two dimensional Bin Packing algorithms. Our second aim is to simulate and compare gathered results based on performance of these algorithms and their applicability on industrial world.

The structure of the paper is as follows: Section 2 discusses interconnection between the user request dimensions in OFDMA and Bin packing algorithms, while section 3 presents a thorough analyzes of developed algorithms, especially Guillotine First Fit and Shelf First Fit algorithms. Section 4 presents the simulation results and critically discusses usage of each algorithm in industry. Finally, section 5 presents the overall conclusions.

¹ Radio access part of telecommunication networks, presents an intermediate communication part that enables the data processing (encoding, multiplexing and transmitting user requests) between the User Equipments and the Core Network part of the system

² Core part of the network presents the main functional interconnection point which routes the user requests to the provider (example: requests for telephone calls to PSTN)

³ Digital data transmission unit, containing encoded user request and additional synchronization bits

2 Affiliation of User Request Dimensions in OFDMA and Bin Packing Algorithms

Two-dimensional Bin packing problems deal with allocation of a set of n rectangles each having width (w) and height (h), into a predefined number of bins, with corresponding dimensions (W, H), without overlapping [15-18].

Packing problems are related to arranging an amount of items into predefined available resources (Fig. 1). In this type of problems a number of elements with different sizes have to be packed into a finite number of bins with different size and forms [19-21].

The objective is to minimize the number of used bins. These algorithms may be used to arrange requests which are multiplexed using OFDMA technique.

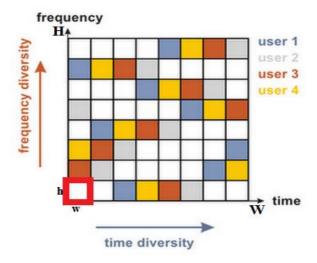


Fig. 1. OFDMA multiplexing schema [22]

OFDMA is a technique that enables all users to transmit and receive at the same time even within single channel on what are called subchannels [23, 24]. In these systems time and frequency are used to separate multiple user signals (Fig. 1.). These dimensions are used in algorithms as input values for width and height of one rectangle which needs to be 'packed' into a specific shelf/bin. We should take into consideration that the rectangles present user requests; shelf and bin dimensions present the predefined dimensions of frames that are going to be used in order to arrange user requests. These terms (rectangle, bin, or shelf) are used just for adoption with terms of Bin-packing algorithms.

The interconnection of two dimensional packing algorithms and OFDMA multiplexing structure are found also in similar approaches. According to [12] a simple heuristic algorithm for the two-dimensional rectangular mapping may be used for downlink bursts in IEEE 802.16e Mobile WiMAX, ensuring strict QoS requirements. This algorithms name is eOCSA (enhanced-One Column Striping with non-increasing

Area first mapping) and it presents an approach that uses the rectangle mapping problem. This algorithm first maps the resources for each subscriber into a downlink burst in a rectangular type. There is also a previous version which maps the resource allocation from bottom to top and right to left [13]. The eOCSA aims to maximize the throughput through sorting the resource allocations in a descending order (largest first). Then, the allocations are mapped from bottom to top and right to left to allow the space for the variable.

Aiming to reduce the waste of utilizing bandwidth, a group of researches in [14] proposed a method that dynamically adjusts the downlink-to-uplink ratio. Dynamic Ratio Determination (DRD) monitors the mapping operation of both downlink and uplink sub-frames, and while considering load balance it receives feedback from both processes and proceeds to the appropriate selection of the forthcoming downlink-to-uplink ratio.

Bin packing algorithms may consider different scaling factors which imply usage of different properties for arrangement. In [25] the following most common algorithms are described:

- Shelf algorithms a shelf is defined to be a subrectangle of the bin with width Wb and height Hs, and as packing proceeds the rectangles are placed into shelves, bottom-up, and from left to right
- Guillotineable algorithms a procedure of placing a rectangle to a corner of a free rectangle of the bin, after which the remaining L-shaped free space is split again into two disjoint free rectangles

Regardless the used algorithm, the primary aim is to map user requirements, which from algorithms' point of view are rectangles.

3 Developed Bin Packing Algorithms

Considering various Bin packing, we have tried to analyze and design ones that are applicable to our practical aim. As we intend to provide an effective bandwidth utilization solution to a dynamic system (telecommunication system which has unpredictable user requests), we chose Shelf and Guillotine algorithms. The development process of these algorithms is based on Software Engineering methodologies.

Shelf Algorithms present the simplest methods to pack rectangles. In these algorithms a shelf is defined to be a subrectangle of the bin with width (Wb) and height (Hs). The packing area is arranged with placement of rectangles into shelves from bottom-up, and from left to right. Hence, when packing a rectangle into the shelf we have to choose whether to rotate it or not. The orientation procedure is applied to all Shelf Algorithms. This procedure considers the suitable way in which the rectangle (respectively the user request) will fit into the allocated area and arrange them so as to save the predefined space that can be used. The orientation includes steps were the rectangle should be placed upright: $R = (\min (w, h), \max (w, h));$ or sideways: $R = (\max (w, h), \min (w, h))$.

Shelf Next Fit algorithm is one of the Shelf Algorithms that helps to solve the packing problem. This algorithm first determines the proper orientation of the rectangle, and then it tries to fit it in the current open shelf. If the rectangle does not fit there, this algorithm advises to open a new shelf (meanwhile closing the current opened shelf), if there is room for that; and if not it terminates execution. A closed shelf will not be opened again.

Shelf Best Width Fit Algorithm considers the shelf in which the remaining capacity or width of the shelf space while placing the incoming rectangle is minimized. This algorithm is very similar to Shelf First Fit Algorithm which we will present in more detail in next subsection.

3.1 Shelf First Fit Algorithm

Shelf Next Fit Algorithm does not uses the empty left areas while closing one shelf and opening a new one. Therefore, it is considered wasteful to leave behind free spaces in old shelves [16, 18, 25, 26]. Shelf First Fit algorithm tries to search into those left areas, which in telecommunication point of view are idle slots, and fit the incoming rectangle/user request there.

What differentiates our approach from the proposed working methodology of Shelf First Fit algorithm in [25]; is that we first try to fit the incoming user request into the current open shelf and if it does not fit there we search into the left areas starting from the lowest indexes, as shown in Fig. 2. This methodology promotes a new version of Shelf First Fit algorithm.

Fig. 2 presents the case scenario of having one bin and four incoming user requests. The first three user requests are mapped in the first opened shelf. The fourth incoming user request cannot fit into the left area after placing the third incoming request. Therefore, the Shelf First Fit algorithm advises to search into the left areas, starting from the beginning, As Fig. 2 presents, our fourth incoming user request may fit only above the left area after placing the third user request (indicating the filled arrow). After closing current shelf (first one), its height will be the largest value of all mapped rectangle's heights (in our case, second mapped request's height).

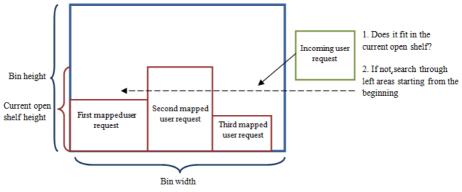


Fig. 2. Shelf First Fit algorithm's operating logic

3.2 Guillotine First Fit Algorithm

Guillotine Algorithms are based on the operation of guillotine split placement. The main idea here is to place an incoming rectangle in the corner of a free-picked rectangle of the bin. After this operation, the guillotine split procedure is applied and the L-shaped free space is composed of two disjoint free rectangles which can be further subdivided. In this way we need to maintain a list of these composed rectangles which can be used to arrange other incoming rectangles.

Split procedure of Guillotine algorithms implies the possibility to choose between two possible directions: horizontal split, if (w<h); otherwise vertical split (w>h). Therefore, while observing rectangle's dimensions we choose which one of the split directions to apply. The split procedure is applied to incoming rectangles in order to place them in the most resourceful position possible. This procedure is applied every time we arrange a rectangle; therefore we can have combined split axes.

The operating logic of Guillotine Algorithm is presented in Fig. 3. The W and H (indicating Width and Height) represent Bin dimensions, while w and h present rectangle dimensions. Two composed rectangles from split procedure are presented as F and F'. Composed rectangles will have dimensions as deduction of bin dimensions with respective rectangle, as shown in Fig. 3. The auxiliary variables w(F) and h(F), are used to store the composed rectangle's width and height dimensions. The following example is related to expressions in Fig. 3:

Bin (12, 15) Incoming rectangle (4, 5)

<u>Vertical Split</u> w (F)= (w) ; h(F)= (H-h) → F (4,10) w (F')= (W-w) ; h(F')= (H) → F' (8, 15)

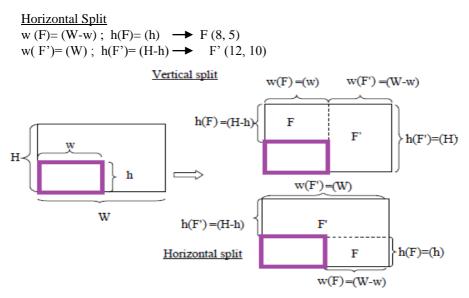


Fig. 3. Split choices used in Guillotine algorithms

First Fit method combined with Guillotine Algorithm's logic can provide a new algorithm which is used to arrange incoming user requests (Fig. 4). We propose Guillotine First Fit Algorithm which takes into consideration all free spaces through split rectangles and places the incoming rectangle into one of them, making the choice based on the numbering (as the name implies-first fit).

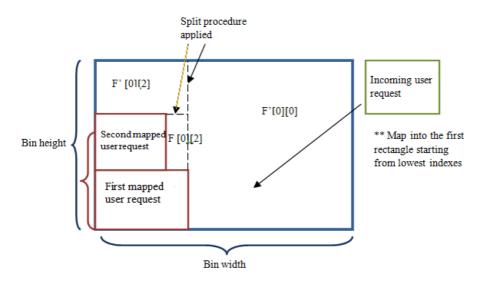


Fig. 4. Guillotine First Fit algorithm's operating logic

During Design and Implementation processes, we have used two dimensional arrays (F [][] and F '[][]) to store the dimensions of rectangles. As it is shown in Fig. 4, while placing the first rectangle (indicated as First mapped user request), the choice for Vertical split is made based on the condition w>h. In this way we will have two free composed rectangles F [0][0] and F'[0][0]. Continuing with second rectangle, we need to search into the left areas (composed rectangles F and F'), starting from the beginning (position [0][0]). After placing the second rectangle (Second mapped user request), we further subdivide the composed F [0][0] rectangle, horizontally (first arrow, denoted with dashed orange color) into another two disjoint rectangles F[0][2] and F'[0][2]. The third incoming rectangle (Incoming user request) can only fit in the composed rectangle F' [0][0].

4 Analysis of Results

In this section we present and discuss results from different scenarios. While simulating these scenarios we provide input values, such as:

- Frame Dimensions
- Number of frames
- User request Dimensions

Provision of user request dimensions is implemented using three different methods. Specifically, they are set according to:

- Fixed value distribution
- Uniform distribution
- Poisson distribution

Following case scenarios will demonstrate how changes of input values and their distribution, may affect the procedure of mapping different user requests with the investigated methodologies, affecting in this way the performance of specific algorithms.

4.1 Fixed Value Distribution

Fixed Value Distribution needs the provision of two dimensions. These are: the user request dimensions of width and height. Final results, for all methods, include number of arranged user requests into predefined frame dimensions and number of idle slots.

Analyzing final results while using fixed distribution of values through our algorithms, we can notice the same performance of all Shelf Algorithms and an effective performance of Guillotine First Fit Algorithm (based on the number of arranged user requests). In Fig. 5 the corresponding numbers of mapped slots and user requests for each algorithm are presented. The number of mapped slots is presented as the sum while multiplying all arranged rectangle's dimensions ((6x9) + (6x9)...). Meanwhile, mapped user requests are calculated as sum of all arranged rectangles.

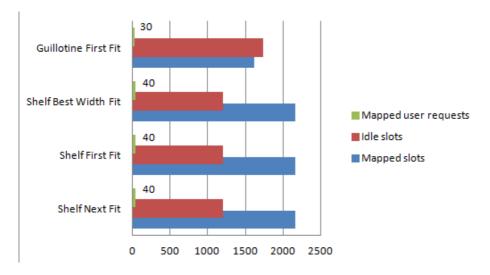


Fig. 5. Simulation results while using Fixed Value Distribution

Shelf Algorithms, while using Fixed Value distribution, will have the same behavior. This due to the fact of the same user request dimensions. This is the case also while increasing the number of used frames, or user request dimensions. The performance of Guillotine First Fit algorithm while using same input user request dimensions is considered effective. Therefore, the number of mapped user requests for GFF algorithm is 30 arranged rectangles, which also presents a higher number of idle slots compared to other algorithms. This due to the incompatible values of composed rectangles (F, F') with user request input values (6, 9), while using shelves with corresponding dimensions (12, 14).

Fixed value distribution is used while simulating this application only for testing reasons. The reason is that: in telecommunication networks there are hardly any cases where all user requests have the same dimensions.

4.2 Uniform Distribution

Aimed to provide different forms of generating input values for our user requests, we implemented the Random number generator. Therefore, a maximum number as a range value for user request's dimensions, is pre-defined. Hence, if the simulator user provides number 9, user request dimensions will have values range from 0 to 8. However, while simulating, we always add 1 in order to avoid user requests of null dimensions.

While generating randomly user request dimensions, we try to compare the performance of our algorithms on the same input data. From the gathered results, we can notice a quite high performance (considered from higher number of mapped slots) of Shelf Next Fit (29%), Shelf First Fit (40%) and Guillotine First Fit algorithms (31%). The Shelf First Fit algorithm exhibits the best performance compared to the other algorithms by managing to arrange 40% of the total number of user slots.

Special effort is put while analyzing the Guillotine First Fit algorithms' performance. Based on this algorithm's operating logic, as the number of frames is increased, the chances to arrange more user requests are higher. This is presented by the linear curve created from gathered results, shown in Fig. 6.

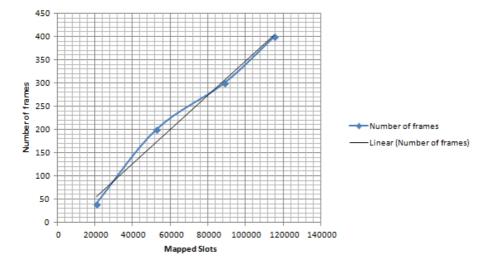


Fig. 6. Uniform Distribution for different number of frames while using Guillotine First Fit Algorithm

In Fig. 6 we present simulation results while considering the same Bin dimensions and changing the number of frames (40, 200, 300, 400) used to arrange randomly generated user requests. We can notice that the performance of GFF is improved, in the context of arranging more user requests, while increasing the number of available slots. The number of arranged user requests (presented in figure as mapped slots), is again calculated as the sum of the multiplied arranged rectangle's dimensions.

4.3 Poisson Distribution

The third method of inputting user request dimensions is based on the Poisson value generator. This method requires two mean values; one for the width of the user request and the other for the height. In this way, the generated user requests have dimensions which range close to the provided mean values.

Analyzing the final results when using the Poisson distribution for generating user request dimensions, we can notice a good performance of all our algorithms. However, in this scenario Shelf Best Width Fit algorithm's efficiency is highlighted.

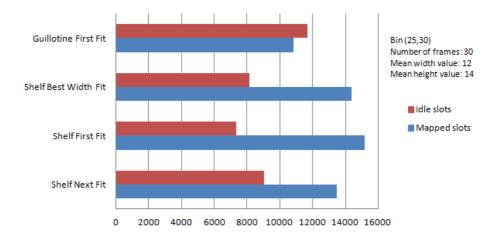


Fig. 7. Performance of algorithms using Poisson distribution

Fig. 7 presents the gathered results as total number of mapped slots and idle slots for each algorithm. Similarly to the previous scenarios, the Shelf First Fit algorithm managed to arrange the highest number of user requests while generating the lowest number of idle slots. Nevertheless, the Shelf Best Width Fit algorithm is very close to the best performance, too. It is reminded that Shelf Best Width Fit uses the methodology of arranging the incoming user request into the left position where the remaining capacity is the lowest. Therefore, as we can notice from the corresponding graphs, the number of idle slots is lower compared to the Shelf Next Fit algorithm.

Guillotine First Fit algorithm performance compared to Shelf First Fit algorithm is considered proximate in the context of mapped user requests. Even though the total number of mapped and idle slots differs, the memory usage and simulation time in the case of Shelf First Fit algorithm are higher. The memory usage for Shelf First Fit and Guillotine First Fit algorithms, while using the same input data and the same allocated memory (32,047,104 B), is presented in Fig. 8. As the figure shows, we should consider a tradeoff between delay, memory usage and efficiency of these algorithms.

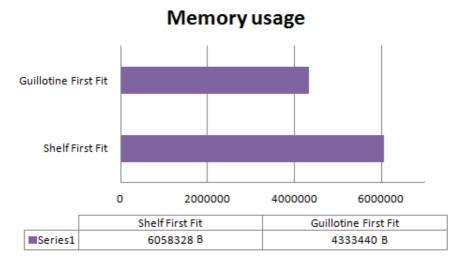


Fig. 8. Memory usage in Shelf First Fit and Guillotine First Fit algorithms

5 Conclusions

The resource allocation problem in the telecommunications area, especially in wireless broadband networks, needs to be carefully addressed. Towards improving the efficiency of the bandwidth utilization, we present different Bin-Packing algorithms which are designed for use in OFDMA networks.

Proposed Bin-Packing algorithms impose different features and functionalities. This helps to compare and distinguish them for specific case scenarios while choosing the distribution method of the user requests.

A new version of Bin Packing algorithms is presented, named: Shelf First Fit and a new developed algorithm named Guillotine First Fit algorithm. The analyses of the results reveal that in the majority of scenarios the performance and efficiency of these algorithms in mapping the incoming user requests, is high. Even though, Shelf First Fit algorithm's operating logic managed that in every performed scenario to be the most resourceful, the consumed memory and simulation time should be taken into account. However, in cases of higher frame dimensions and number of frames, we should consider Guillotine First Fit algorithm.

Further enhancements of our algorithms for considering QoS requirements are planned. Moreover, the addition of more algorithms (Skyline algorithm, Maximal Rectangle algorithm) for ensuring better bandwidth utilization can provide a better generalized application for wide industrial usage.

References

- Bannister, J., Mather, P., Coope, S.: Convergence Technologies for 3G Networks IP, UMTS, EGPRS and ATM. In: John Wiley & Sons Ltd The Atrium. Southern Gate Chichester. West Sussex (ed) Background to convergence 1rt end. PO19 8SQ, England, ISBN 0-470-86091-X (2004)
- Romero, J.P., Sallent, O., Augusti, R., Diaz-Guerra, M.A.: Radio resource management strategies in UMTS. In: John Wiley & Sons Ltd. The Atrium Southern Gate Chichester. West Sussex (ed) The mobile communications sector 1rt edn. PO19 8SQ, England. ISBN-10: 0470022779 (2005)
- Junaid, M.A., Farooq, A., Shah, A.: Evolution and Development towards 4th Generation (4G) Mobile Communication Systems. Journal of American Science. Volume 6. pp 63-68. ISSN: 1545-1003 (2010)
- Tanaka, K.: The Mobile Communications Environment in the European Union: Systems, Regulations, and Consequences in the Past, Present, and Future. In: IEEE 2001 Conference on the History of Telecommunications. St. John's Newfoundland Canada. Pp 25-27. July 2001
- Marjanovic, D., Müller, C., Schönenberger, T.: New Opportunities for B2B and B2C Services Offered by New Mobile Communication Technologies. Presentation Lectures: Department of Informatics Communication Systems Group. University of Zurich
- 6. Tellabs: 4G, the What, Why and When. The worldwide adoption and growth of wireless are the fastest technological achievements in history. In: White paper, Innovative solutions at your fingertips. 74.1948E Rev. C 2/12 (2012)
- 7. Borcoci, E.: WiMAX Technologies: Architectures, Protocols, Resource Management and Applications. In: CTRQ Conference. Bucharest, Romania, June 29 July 5, 2008
- So-In, Ch., Jain, R., Tamimi, A.-K.: Scheduling in IEEE 802.16e Mobile WiMAX Networks: Key Issues and a Survey. Vol27, No2, February 2009
- 9. Hwang, J.S., Consulta, R.R., Youn, H.J.: Opportunities & Threats for implementation of 4G cellular networks. Paper published in online library at http://www.scribd.com/doc/6852728/4G, PTC'07 Proceedings (2012)
- Rumney, M.: IMT-Advanced: 4G Wireless Takes Shape in an Olympic Year. In: Agilent Measurement Journal. USA. 5989-9793EN (2008)
- 11. Omerovic, S.: WiMax Overview. Faculty of Electrical Engineering. Dissertation, University of Ljubljana, Slovenia
- So-In, C., Jain, R., Al-Tamimi, A.: eOCSA: An Algorithm for Burst Mapping with Strict QoS requirements in IEEE 802.16e Mobile WiMAX Network. In: 2nd Inter. Conf. on Comp. and Automation Engineering (ICCAE 2010)
- So-In, C., Jain, R., Al-Tamimi, A.: OCSA: An algorithm for Burst Mapping in IEEE 802.16e Mobile WiMAX Networks. In: 15th Asia-Pacific Conference on Comm. (APCC 2009)
- Sarigiannidis, A., Nicopolitidis, P., Papadimitriou, G., Sarigiannidis, P., Louta, M.: Using Learning Automata for Adaptively Adjusting the Downlink-to-Uplink Ratio in IEEE 802.16e Wireless Networks. In: ISCC '11 Proceedings of the 2011 IEEE Symposium on Computers and Communications, Pages 353-358. ISBN: 978-1-4577-0680-6. doi 10.1109/ISCC.2011.5983863 (2011)
- Lodi, A., Martello, S., Vigo, D.: Recent advances on two-dimensional Bin packing problems. In: Elsiver. Volume 123. Issues 1–3. Pages 379–396. PII: S0166-218X(01)00347-X (2002)

- Zhang, G.: A new version of on-line variable-sized Bin packing. In : Elseiver. Volume 72. Issue 3. Pages 193–197. AMS classification: 90B35, 9OC27 (1997)
- 17. Bin Packing. Presentation, The Icon Programming Language: The University of Arizona, http://www.cs.arizona.edu/icon/oddsends/bpack/bpack.htm
- Dosa, G., Sgall, J.: First Fit packing: A tight analysis. In: Leibniz International Proceedings in Informatics Schloss Dagstuhl. pp. 1–15. Doi: 10.4230/LIPIcs.xxx.yyy.p (1998)
- Kühn, P.J.: Designing a Frequency Selective Scheduler for WiMAX using Genetic Algorithms. Dissertation, University of Stuttgart, Institute of Communication Networks and Computer Engineering (2008)
- 20. Davis, T.: Bin Packing, http://www.geometer.org/mathcircles (2006)
- 21. Lodi, A.: Algorithms for Two-Dimensional Bin Packing and Assignment Problems. Doctorate thesis, Universita Degle Studi Di Bologna (1999)
- 22. History of OFDMA and how it Works (2005) http://www.webopedia.com/DidYouKnow/Computer_Science/2005/OFDMA.asp
- 23. Shazzad, K.M.: Orthogonal frequency-division multiplexing, http://www.mikrobiokosmos2009.org/7-orthogonal-frequency-division-multiple-accessofdma.html
- 24. Ahmadi, S.: Introduction to mobile WiMAX Radio Access Technology: PHY and MAC Architecture. Presentation: Wireless Standards and Technology Intel Corporation (2006)
- 25. Jylänki, J.: A Thousand Ways to Pack the Bin A Practical Approach to Two-Dimensional Rectangle Bin Packing. Dissertation. Available at http://clb.demon.fi/files/RectangleBinPack.pdf, 2010 (2010)
- Bin packing. Chapter 10. In: Institute of Computer Science, Faculty of Mathematics and Natural Sciences II, part of the Lecture notes http://www2.informatik.hu-berlin.de/alkox/lehre/lvws1011/coalg/bin_packing.pdf

On the Role of Stream Reasoning in Run-time Monitoring and Analysis in Autonomic Systems

Rustem Dautov¹, Mike Stannett² and Iraklis Paraskakis¹

¹ South-East European Research Centre (SEERC), City College – International Faculty of the University of Sheffield, 24 Proxenou Koromila Street, 54622 Thessaloniki, Greece {rdautov, iparaskakis}@seerc.org

² Department of Computer Science, University of Sheffield, Regent Court, 211 Portobello Street, Sheffield S1 4DP, United Kingdom m.stannett@dcs.shef.ac.uk

Abstract. Both academia and industry have been putting a lot of effort into investigating various solutions to the problem of autonomic management of complex distributed computer systems. Existing approaches, based on the principles of autonomic computing, typically rely on monitoring a managed system, analysing the monitored values, and executing corresponding adaptation actions. By addressing existing limitations, we present a novel research area – stream reasoning – in the context of run-time monitoring and analysis in autonomic systems. This research area aims at enhancing existing stream processing techniques by adding dynamic reasoning support, thereby fitting scenarios when timely, yet precise analysis of dynamically flowing data is needed. In this paper, by discussing benefits and shortcomings, we speculate on the potential role of the approach in autonomic computer systems.

1 Introduction

The increasing complexity, cost and heterogeneity of distributed computer systems have motivated researchers to focus on the development of self-management mechanisms following the principles of autonomic computing. The goal of autonomic computing is to create computer systems capable of managing themselves to a far greater extent than they do today [1]. The fundamental requirements for a self-managing system to be effective are self- and context-awareness: in order to adapt to changes in its operating environment, a computer system should first observe these changes and then correctly interpret them – that is, to support self-management, it should perform both self-monitoring and contextual analysis.

Even though the interpretation of monitored values is the key to prompt and efficient adaptations, the analysis support in existing approaches aiming at implementing self-managing systems, is still hindered with certain limitations [2] – analysis rules are hard-wired into the code, analysis of the monitored values is delayed, it is not well automated, etc. To address (at least partially) these constraints, in this paper we describe the potential of *stream reasoning* – a novel research area at the intersection of the Semantic Web and stream processing – to support run-time monitoring and analysis. We present a non-exhaustive list of benefits of utilising stream reasoning techniques when implementing monitoring and analysis mechanisms within autonomic systems, including, e.g., timeliness, intelligence, separation of concerns, and higher automation. The ideas to be presented have potential shortcomings as well, which are also discussed in this paper. However, positive aspects of this approach make us believe that they are worth further investigation.

The rest of the paper is organised as follows. Section 2 gives a short introduction to the motivation behind the presented work – the need for enhanced run-time monitoring and analysis in distributed computer systems. It lists some challenges associated with monitoring such complex systems, and outlines some basic principles an effective analysis mechanism should follow. Section 3 addresses these challenges and principles by introducing the notion of data streams. It also briefly explains existing approaches in this field, such as data stream management systems and Complex Event Processing. Section 4 presents the concept of stream reasoning as a Semantic Webenabled extension to stream processing – it lists potential benefits stream reasoning can bring in the context of run-time monitoring and analysis. Section 5 discusses some potential shortcomings associated with the presented ideas, and Section 6 concludes the paper.

2 Background and Motivation

2.1 Need for Monitoring

Due to the growing importance of distributed systems such as service-based applications, clouds and grids in recent years, the scientific community focused on the theoretic foundations and research on how to make such complex systems adaptive and sustainable, often referring to the original self-* principles of autonomic computing [3, 4] – a concept that brings together many fields of computing with the purpose of creating computer systems that self-manage. IBM was the first to introduce this term in 2001 [5] to refer to systems which are able to adapt themselves to changes happening in the surrounding environment, and suggested a reference model for creating closed adaptation loops known as MAPE-K (Monitor, Analyse, Plan, and Execute based on Knowledge). IBM compared complex computer systems to the human body, and suggested that such systems should also demonstrate certain autonomic properties, that is, should be able to independently take care of regular maintenance and optimization tasks, thus reducing the workload on system administrators [6]. These challenges of creating closed control loops had been already addressed in other research areas, including both Agent Theory [6, 7] and Control Theory [6], which inspired IBM in their vision of autonomic computer systems.

Inspired by IBM's Autonomic Computing manifesto and referring to the original self-* characteristics of autonomic systems, early theoretical works served as a groundwork for more and more prototypical implementations of self-managing mech-

anisms in various computer systems [3]. Existing self-adaptation mechanisms typically follow the agent-driven approach of sensing and acting upon sensed values by (fully or partially) implementing control feedback loops, such as the already mentioned MAPE-K model or the CADA model (Collect, Analyse, Detect, Act) [8]. The common component in most of the solutions implementing autonomic computing principles is a monitoring mechanism which is the "trigger" either for further run-time adaptations, or else for post-mortem data analysis. In a broad sense, monitoring may be defined as a process of collecting and reporting relevant information about the execution and evolution of a computer system [9].

Depending on a particular purpose of a designed service-based system, monitoring activities can be used [3]:

- to perform run-time analysis of system correctness, i.e., checking the system execution and its parameters against certain specifications that distinguish "correct" and "failing" situations;
- to perform the diagnosis and recovery of any identified faults;
- to instantiate and guide the optimisation activities regarding resource allocation (e.g., in a cloud computing scenario);
- to support dynamic adaptation of the application to the relevant change in its environment;
- to support the long-term adaptation (i.e., evolution) of a computer system, etc.

All these monitoring processes target the collection of data for a specific artefact, referred to as the *monitored subject* [9]. Depending on the context, the range of monitored subjects can be rather broad, including applications, operating systems, networks, Web servers, etc.

Two types of monitoring are usually identified in the literature [6]:

- *Passive* monitoring, also known as *non-intrusive*, assumes that no changes are made to the managed element. This kind of monitoring is generally targeted at the context of the managed element.
- *Active* monitoring, also known as *intrusive*, entails designing and implementing software in such a way that it provides some entry-points for capturing required properties (e.g., APIs).

Usually, monitoring is performed under one or both of the following main modes for collecting information [9]:

- *Polling mode*: querying or directly interacting with the monitored subject within regular time intervals.
- *Push mode*: propagating events or other data from the information sources of the monitored subject to the monitoring mechanism.

Even though monitoring activities also include such techniques as post-mortem log analysis, data mining, online and offline testing, etc. [10], throughout the rest of the paper we will only refer to the "classical" notion of run-time monitoring.

2.2 Need for Analysis

Monitoring on its own is not enough to enable adaptations to be carried out. For example, suppose we are monitoring response times from a Web service, and the observed value is 5 seconds. What does this value tell us? How can we know if it is a slow response or not? What is really needed is the analysis support, so that comparing observed values against expected ones, we can decide whether the current state of the system is faulty or potentially leads to a failure. The analysis component's main responsibility is to assess the current situation and detect failures or suboptimal behaviours of the managed system. In other words, when coupled with an analysis component, monitoring will be important in supporting problem determination and recovery when a fault is found or suspected [11]. Some of the common drawbacks existing analysis approaches tend to suffer from, are [2, 10]:

- *Rigidness*: analysis rules are "hard-coded", which hinders possible modifications.
- *Low intelligence*: analysis mechanisms are not sophisticated and intelligent enough to produce a precise diagnosis when it comes to complex, non-trivial scenarios.
- *Delayed execution*: analysis processes take place after monitored values have been observed, which decreases the actuality of the results.
- *Low level of automation*: the analysis is (fully or partially) performed by human operators, which again results in low actuality of the results and human errors.

With the development of highly distributed, complex and dynamic systems, such as service-based application systems, clouds, grids, etc. where large volumes of data are continuously generated and consumed, timely problem determination becomes an even more challenging task. In this respect, from an information management point of view, we can distil the following challenges of existing computer systems [12], which have to be taken into consideration when implementing monitoring and analysis mechanisms:

- *Dynamism*: various sources of information are constantly generating data (which is then processed, stored, deleted, etc.) at an unpredictable rate. Moreover, various system components are always evolving, so that new sources of information are coming up, while old ones are disappearing, thus making the whole system even more dynamic.
- *Distributed nature*: the information may come from various logically and physically distributed sources. The first means that it may originate from databases, file systems, running applications, or external Web services. The latter refers to the fact that all these "logical" sources may be deployed in separate virtual machines, servers and even distinct datacentres.
- *Large volumes*: the amount of raw data being generated (for example, within a social network) is huge. Even if we neglect those information flows that are not directly relevant to the context of self-management (i.e., the so-called "noise"), the amount of data remaining is still considerable.
- *Heterogeneity*: originating from various distributed sources such as applications, databases, user requests, external services, etc., the information is a priori hetero-

geneous. Apart from the heterogeneity in the data representation, such as differences in data formats/encodings, there is also heterogeneity in the semantics of the data. For example, two completely separate applications from different domains with different business logic may store logging data in XML. In this case, the data is homogeneous in its format and, potentially, structure, but completely heterogeneous at the semantic level.

Accordingly, in order to address these challenges, the analysis mechanism should follow the following principles:

- *Automation*: obviously, with the ever-expanding complexity of software systems, the process of analysing monitored data should be automated. The automated analysis may be complemented by human supervision, but the main process should be performed in an autonomic manner.
- *Run-time execution*: the analysis should be performed dynamically that is, as soon as new data arrives it has to be evaluated. As opposed to static approaches, where monitored values are first stored (e.g., in a database or log files) for later analysis, dynamic analysis assumes we will be processing streams of incoming data "on the fly", without storing it on a hard drive.
- *Flexibility:* the analysis mechanism should be flexible enough so as to accommodate the evolving and expanding nature of existing software systems. Consider a scenario where the analysis logic is "hardcoded" with numerous "if-then" and "switch-case" operators – as opposed to declarative approaches to defining rules, introduction of a new rule to the analysis mechanism in this circumstances will require rewriting the code, recompiling, and redeploying/restarting the whole system [13].

3 Data Streams

The challenges associated with run-time monitoring and analysis of contemporary complex distributed software systems, listed in the previous section, are already attracting researchers' attention. In the era of Big Data, we create 2.5 quintillion bytes of data every day – so much that 90% of the data in the world today has been created in the last two years alone [14]. This data comes from everywhere: sensors used to gather climate information, posts to social media sites, digital pictures and videos, purchase transaction records, and cell phone GPS signals to name a few. Even though existing technologies seem to succeed in storing these overwhelming amounts of data, on-the-fly processing of newly generated data is a challenging task.

An increasing number of distributed applications are required to process continuously streamed data from geographically distributed sources at unpredictable rates to obtain timely responses to complex queries [15]. A key research area addressing these issues is Information Flow Processing (IFP); in contrast to the traditional static processing of data, IFP focuses on *flow processing* and *timeliness* [16]. The former means that data is not stored, but rather continuously flowing and being processed, and the latter refers to the fact that time constraints are crucial for IFP systems. These

two main requirements have led to the emergence of a number of systems specifically designed to process incoming information streams according to a set of pre-deployed processing rules. In this context, a *data stream* comprises an unbounded sequence of continuously appended values, each of which carries a timestamp that typically indicates when it has been produced [12, 15]. Usually (but not necessarily), recent values are more relevant and useful, because most applications are interested in processing current observations to achieve near real-time operation. Examples of data streams include sensor values, stock market tickers, social status updates, heartbeat rates, etc. To cope with the unbounded nature of streams and temporal constraints, so-called continuous query languages [15] have been developed to extend the conventional SQL semantics with the notion of windows. A window transforms unbounded sequences of values into bounded ones, allowing the traditional relational operators to be applied. This approach restricts querying to a specific window of concern which consists of a subset of statements recently observed on the stream, while older information is (usually) ignored [12, 17]. Windows can be specified both in terms of tu*ples*, where a window comprises a given number of most recent elements regardless of arrival time; and *time*, where a window comprises all elements which have arrived during some specified time frame.

The concepts of unbounded data streams and windows are visualised in Figure 1. The small squares represent tuples continuously arriving over time and constituting a data stream, whereas the thick rectangular frame illustrates the window operator applied to this unbounded sequence of tuples [12]. As time passes and new values are appended to the data stream, old values are pushed out of the specified window, i.e. they become irrelevant and may be discarded (unless there is a need for storing historical data for later analysis).

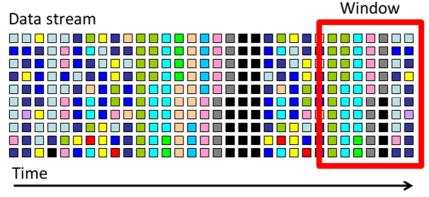


Fig. 1. A data stream and a window (modified from [17]).

Data stream management systems (DSMSs) – an evolution from traditional static database management systems – have been developed to process streams of data coming from different sources to produce new data streams as output [16]. They target transient, continuously updated data and run standing (i.e., continuous) queries, which fetch updated results as new data arrives. Complex Event Processing (CEP) goes one step further and tries to detect complex event patterns, themselves consisting of sim-

pler atomic events, within a data stream [16]. Accordingly, from this point of view, flowing information items can be seen as notifications of events happening in the external world. An event may be anything that changes the current state of affairs, e.g. a fire alarm, social status update, change of traffic light, etc. Accordingly, the focus of this perspective is on detecting occurrences of particular patterns of (low-level) events that represent higher-level events. A notification of a complex event to the interested parties is triggered if and only if a corresponding pattern of lower-level events is detected.

4 Stream Reasoning

Stream Reasoning research [18] goes even further by enhancing the continuous querying with the run-time reasoning support – that is, with capabilities to infer additional, implicit knowledge based on already given, explicit facts. Barbieri et al. [21], who were the first to use the term 'stream reasoning' for the new concept, defined it as "reasoning in real time on huge and possibly noisy data streams, to support a large number of concurrent decision processes". As data streams have become more and more common on the Web (e.g. stock exchange movements, weather information, sensor readings, social networking activity notifications, etc.), the combination of data stream processing techniques with data streams distributed across the Web came as a natural fit, and this in turn required new ways of coping with the typical openness and heterogeneity of the Web environment – in this context, the role of Semantic Web technologies is to facilitate data integration in open environments, and thus help to overcome these problems by using uniform machine-readable descriptions to resolve heterogeneities across multiple data streams [19]. For example, the Semantic Sensor Web [20] represents an attempt to enable more expressive representation, advanced access, and formal analysis of avalanches of sensor values in such domains as traffic surveillance, environmental monitoring, house automation and tracking systems, by encoding sensor descriptions and sensor observation data with Semantic Web languages.

As Semantic Web technologies are mainly based on Description Logics, their application to data stream processing also offers new opportunities to perform reasoning tasks over continuously and rapidly changing flows of information. In particular, stream reasoning utilises and benefits from the following Semantic Web technologies:

- *Resource Description Framework*¹ (*RDF*) as a uniform format for representing streamed heterogeneous data as a collection of (*subject, predicate, object*) triples using a vocabulary defined in an OWL ontology;
- *OWL* ontologies² and *SWRL* rules³ as a source of static background knowledge, which may also act as a vocabulary of terms for defining RDF triples;

¹ http://www.w3.org/RDF/

² http://www.w3.org/TR/owl-features/

³ http://www.w3.org/Submission/SWRL/

• *SPARQL*-based⁴ continuous query languages as a way of querying RDF streams and performing reasoning tasks by combining them with the static background knowledge.

As a result, several prominent stream reasoning approaches have emerged: C-SPARQL [22], CQELS [23], ETALIS [24], and SPARQL_{stream} [15] to name a few. These stream reasoning systems aim at preserving the core value of data stream processing, i.e. processing streamed data in a timely fashion, while providing a number of additional features [19]:

- Support for advanced reasoning: depending on the extent to which stream reasoning systems support reasoning, it is possible not only to detect patterns of events (as CEP already does), but in addition to perform more sophisticated and intelligent detection of failures by inferring implicit knowledge based on pre-defined facts and rules (i.e., static background knowledge).
- Integration of static background knowledge with streamed data: it is possible to match data stream values against a static background knowledge base (usually represented as an ontology), containing various facts and rules. This separation of concerns allows for seamless and transparent modification of analysis rules constituting the static knowledge base.
- Support for expressive queries and complex schemas: ontologies, acting as static background knowledge, also serve as a common vocabulary for defining complex queries. This means that classes and properties constituting an ontology provide "building blocks" and may be used for defining queries of required expressivity.
- Support for logical, data and temporal operators: to cope with the unlimited nature of data streams, stream reasoning systems extend conventional SQL-based logical and data operators with temporal operators. This allows us to limit an unbounded stream to a specific window, and also to detect events following one after another chronologically.
- Support for time and tuple windows: as described earlier, windows may be specified either by time-frame, or else by the number of entries to be retained, regardless of arrival time.

Taking into consideration these features of stream reasoning systems, we believe that they provide a promising foundation for implementing monitoring mechanisms with built-in analysis support. Our initial experiments implementing monitoring and analysis mechanisms within a cloud application platform using C-SPARQL, an OWL ontology, and SWRL rules, are reported in [12], and suggest that this approach is viable and has the potential to provide the required level of analysis support.

⁴ http://www.w3.org/TR/rdf-sparql-query/

5 Demonstrating Benefits of Stream Reasoning

To demonstrate the benefits of Stream Reasoning techniques in the context of runtime monitoring and analysis of data streams, we now present a simple scenario and describe how it can be potentially implemented using: (i) traditional stream processing, (ii) Complex Event Processing, and, finally, (iii) Stream Reasoning.

Let us assume that we are interested in monitoring response times from various Web services constituting a composite service-based application system, so as to detect, for example, if a web server hosting a group of services has crashed or a particular service is getting overloaded. The examples presented below are intended for illustrative purposes, and are correspondingly simplified.

5.1 Monitoring and Analysis with Traditional Stream Processing

Using traditional stream processing techniques, we are able to register corresponding continuous queries, which would trigger every time a new response time value appears on the data stream. Then these monitored values should be fed to an analysis engine – typically, a hard-coded software component, which implements the analysis routine. For example, to detect if response time from a particular service has been gradually increasing during a given time interval – that is, to predict if it will soon become overloaded, the analysis engine should evaluate input values against a predefined set of hard-coded rules (e.g., numerous "if-then" and "switch-case" operators). Obviously, this approach suffers from complexity, low performance, low flexibility and potential human error.

5.2 Monitoring and Analysis with Complex Event Processing

CEP can help improve the situation by partially shifting the analysis routine to the monitoring component. For example, a single change in response time from a Web service can be seen as an atomic event, whereas a sustained pattern of increasing response times from a given Web service represents a complex event. By employing CEP techniques, we are exempted from the routine of defining temporal relations between monitored events – instead, we use a temporal query operator, and in the WHERE clause of the corresponding continuous query we specify the order of the events to be detected. It is worth noting, however, that defining CEP queries is not something straightforward, and, depending on the complexity of the monitored events, may result in clumsy, multi-level and nested constructions.

5.3 Monitoring and Analysis with Stream Reasoning

As we have seen, stream reasoning allows us to benefit from integrating static background knowledge with data streams, and performing reasoning tasks over values observed on those streams based on this knowledge. Let us assume that we have a background ontology, which, among other things, includes knowledge about which Web services are hosted on which Web servers. Unlike CEP, which only allowed us to detect an increase in response time from a particular Web service, adding in the background knowledge allows us to detect cases where *several* services, all hosted on the same server, get overloaded, indicating a situation where the *Web server itself* is becoming overloaded.

A capability to resolve subclass relationships at run-time is another benefit of stream reasoning in the context of run-time monitoring and analysis. Instead of registering (hundreds of) separate queries for every single instance of the Web services being monitored (as we would do in CEP), we can register a single query for the superclass Service, and then define a simple background taxonomy specifying its subclasses. This taxonomy helps the stream reasoner understand whether an element observed on the stream is a subclass of Service, and should therefore be evaluated against the defined query.

6 Potential Shortcomings

Despite the benefits we have described of using stream reasoning in the context of run-time monitoring and analysis, this approach is not a "silver bullet" and has potential shortcomings. We will try to summarise them in this section:

- *Need for unified data representation*: in order to apply reasoning to a data stream, its heterogeneous values must first be represented in a common format RDF. Only after monitored data is represented in the appropriate RDF can it be queried and reasoned about.
- *No standards yet:* while other technologies constituting the Semantic Web technology stack have been already standardised [25], research effort in the stream reasoning area is still disjoint and scattered. Standardisation by the World Wide Web Consortium⁵ (W3C) is currently under way, but it may take another several years until agreed standards are widely adopted.
- *Immature reasoning support:* as opposed to the reasoning capabilities of SPARQL a standardised static SQL-like query language for RDF data, widely used within the Semantic Web community querying over data streams with reasoning support has not as yet been fully implemented. None of the existing continuous query languages for RDF streams fully supports the SPARQL 2.0 specification.
- Low performance: stream reasoning research is still in its infancy, and the performance of dynamic reasoning over a data stream is an issue. Expressivity of a query language is known to be inversely related to its performance [19] the more expressive queries are, the longer it takes to execute them. This affects both the scalability of stream reasoning systems, and the actuality and accuracy of the results obtained. At the moment, much research effort is concentrated on advancing the reasoning support, while performance issues, such as query optimisation, caching, and parallel querying are yet to be explored.

²⁵⁶

⁵ http://www.w3.org/

7 Conclusion

We have presented run-time monitoring as a foundation for implementing autonomic computer systems. However, in order for the monitored values to be of use in the further adaptation processes, there should be some form of evaluation over these values, so that potential failures of the monitored system can be detected and/or predicted. By listing challenges associated with run-time monitoring of complex distributed software systems, we have distilled principles to be followed when implementing a suitable analysis mechanism. In this context, we introduced stream reasoning - an extension of data stream processing, which allows for run-time reasoning over data streams – as a way of enhancing monitoring mechanisms with analysis capabilities. Stream reasoning techniques facilitate the evaluation of monitored data streams in a timely manner by supporting static background knowledge, expressive continuous query languages and temporary query operators. We have also described some of the potential shortcomings associated with stream reasoning in run-time monitoring and analysis solutions, arising from the relative immaturity of the underlying research areas. However, stream reasoning has attracted increasing attention in recent years, and we can reasonably expect these limitations to be addressed in the near future.

References

- Tesauro, G., Chess, D.M., Walsh, W.E., Das, R., Segal, A., Whalley, I., Kephart, J.O., White, S.R.: A multi-agent systems approach to autonomic computing. Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 1. pp. 464–471 (2004)
- 2. Delgado, N., Gates, A.Q., Roach, S.: A taxonomy and catalog of runtime software-fault monitoring tools. Ieee Trans. Softw. Eng. 30, 859–872 (2004)
- Breskovic, I., Haas, C., Caton, S., Brandic, I.: Towards Self-Awareness in Cloud Markets: A Monitoring Methodology. Ninth IEEE International Conference on Dependable, Autonomic and Secure Computing (DASC) (2011)
- 4. Kephart, J.O., Chess, D.M.: The vision of autonomic computing. Computer. 36, 41–50 (2003)
- Horn, P.: Autonomic Computing: IBM's Perspective on the State of Information Technology. Comput. Syst. 15, 1–40 (2001)
- 6. Huebscher, M.C., McCann, J.A.: A survey of autonomic computing—degrees, models, and applications. Acm Comput Surv. 40, 1–28 (2008).
- Wooldridge, M., Jennings, N.R.: Intelligent agents: Theory and practice. Knowl. Eng. Rev. 10, 115–152 (1995)
- Dobson, S., Denazis, S., Fernández, A., Gaïti, D., Gelenbe, E., Massacci, F., Nixon, P., Saffre, F., Schmidt, N., Zambonelli, F.: A survey of autonomic communications. Acm Trans. Auton. Adapt. Syst. Taas. 1, 223–259 (2006).
- Bratanis, K.: Towards engineering multi-layer monitoring and adaptation of service-based applications. SEERC Technical Reports. Thessaloniki, Greece: South-East European Research Centre (SEERC) (2012)
- Kazhamiakin, R., Benbernou, S., Baresi, L., Plebani, P., Uhlig, M., Barais, O.: Adaptation of Service-Based Systems. In: Papazoglou, M.P., Pohl, K., Parkin,

M., and Metzger, A. (eds.) Service Research Challenges and Solutions for the Future Internet. pp. 117–156. Springer Berlin Heidelberg (2010)

- Kephart, J.O., Chess, D.M.: The vision of autonomic computing. Computer. 36, 41 – 50 (2003)
- Dautov, R., Paraskakis, I., Kourtesis, D., Stannett, M.: Addressing Self-Management in Cloud Platforms: a Semantic Sensor Web Approach. Proceedings of the International Workshop on Hot Topics in Cloud Services (HotTopiCS 2013)., Prague, Czech Republic (2013)
- 13. Dautov, R., Paraskakis, I., Kourtesis, D.: An ontology-driven approach to selfmanagement in cloud application platforms. Proceedings of the 7th South East European Doctoral Student Conference (DSC 2012). Thessaloniki, Greece (2012)
- 14. IBM: Bringing Big Data to the Enterprise, http://www-01.ibm.com/software/data/bigdata/
- 15. Calbimonte, J.-P., Jeung, H., Corcho, O., Aberer, K.: Enabling Query Technologies for the Semantic Sensor Web. Int. J. Semantic Web Inf. Syst. (2012)
- Gucola, G., Margara, A.: Processing Flows of Information: From Data Stream to Complex Event Processing. Acm Comput. Surv. (2011)
- Barbieri, D., Braga, D., Ceri, S., Della Valle, E., Grossniklaus, M.: Stream Reasoning: Where We Got So Far. Proceedings of the 4th International Workshop on New Forms of Reasoning for the Semantic Web: Scalable and Dynamic (NeFoRS) (2010)
- Della Valle, E., Ceri, S., Barbieri, D., Braga, D., Campi, A.: A First Step Towards Stream Reasoning. In: Domingue, J., Fensel, D., and Traverso, P. (eds.) Future Internet. pp. 72–81. Springer Berlin / Heidelberg (2009)
- Lanzanasto, N., Komazec, S., Toma, I.: Reasoning over real time data streams. http://www.envision-project.eu/wpcontent/uploads/2012/11/D4.8-1.0.pdf
- Sheth, A., Henson, C., Sahoo, S.S.: Semantic sensor web. Internet Comput. Ieee. 12, 78–83 (2008)
- Barbieri, D., Braga, D., Ceri, S., Valle, E.D., Huang, Y., Tresp, V., Rettinger, A., Wermser, H.: Deductive and Inductive Stream Reasoning for Semantic Social Media Analytics. Intell. Syst. Ieee. 25, 32 –41 (2010)
- 22. Barbieri, D.F., Braga, D., Ceri, S., Della Valle, E., Grossniklaus, M.: C-SPARQL: SPARQL for continuous querying. Proceedings of the 18th international conference on World wide web. pp. 1061–1062. ACM, New York, NY, USA (2009)
- Le-Phuoc, D., Dao-Tran, M., Parreira, J.X., Hauswirth, M.: A native and adaptive approach for unified processing of linked streams and linked data. Proceedings of the 10th international conference on The semantic web - Volume Part I. pp. 370–388. Springer-Verlag, Berlin, Heidelberg (2011)
- 24. Anicic, D., Rudolph, S., Fodor, P., Stojanovic, N.: Stream reasoning and complex event processing in ETALIS. Semantic Web. (2010)
- 25. Hitzler, P., Krötzsch, M., Rudolph, S.: Foundations of Semantic Web Technologies. CRC Press (2009)

Research Method for Studying Complex Phenomenon in Information Systems

Ehimen Ejodame, Stuart Maguire

Management School, The University of Sheffield, United Kingdom

Abstract. This research corroborates the evidence in the literature of the complex interactions and implications of human factors in Information Technology (IT) implementation. There is somewhat limited understanding about the management of these 'softer' and less detectable components; which have been observed to be the critical determinant in performance enhancement. Researchers suggest an apparent gap between the theory and practice thereof. This study looks at the soft (human) aspect considerations of IT implementation in a public sector. The complexities and interrelations of the soft issues necessitate the choice of systems thinking approach. The aim is to undertake a holistic study of the soft issues to enable an all inclusive management of problem elements.

Keywords: Research Methods, Information Technology, Systems Thinking, Qualitative Research, Human Factors and Complex Systems.

1 Introduction

Empirical results show that several Information System (IS) projects have failed while concentrating on the technical aspects of the projects and neglecting the softer (or human) aspects (Maguire and Redman, 2006; Ojako et al, 2010; and Cheikhrouhou et al, 2011). Research conclusions emphasize that the human or soft aspects of IS implementation are the critical determinant for performance (Olorunniwo and Xi, 2010). There is evidence that these soft aspects exhibit untold complex interactions. The factors are interrelated such that they cannot be studied in isolation. Studying such complex elements in isolation often cause important interactions to be overlooked (Checkland and Sholes, 1990; Willcocks, 1993).

Such complex webs of interrelations are often referred to as 'messy' situations which are the characteristics of most of the modern world situations (Checkland and Scholes, 1990). Scientific methods of enquiry have been somewhat unsuccessful in handling such complex situations, hence, the emergence of the systems thinking method of enquiry (Lewis, 1994). Systems' thinking approach is in tune with the dictum that the solution to the whole of a problem is more than the sum of the solutions of its individual parts. This implies that an attempt to create a solution for different elements of a problem often fails to meet the solution of the larger real world problem. Hence, there is the need for a solution that addresses the problem as a whole

in line with systems thinking (Checkland and Scholes, 1990). This study proposes a unique approach in the application of the principles of systems thinking to understand the human aspect interactions of IT implementation in a public sector organisation in Nigeria.

2 Basis of the Proposed Research

Previous studies have highlighted the intricacies and complexities of the human aspects in IT implementation with evidence that the human aspect is the critical determinant for successful IT implementation (Brooke and Maguire, 1998; Maguire and Redman, 2006). An analysis of the supply chain management (SCM) situation of the proposed case revealed a poor performance with clear evidence for the need to adopt the use of IT to manage the system. This will replace the cumbersome manual paper management efforts currently in practice. Preliminary enquiry of the SCM system in the organization revealed a number of softer complex issues relating to the people and processes involved in the organization. The analysis indicate the need to further explore these soft issues in order to gain insight on how to manage them to enhance IT implementation and SCM functions in the organization. Hence the focus of the study is to understand how these human or soft factors of IT implementation can be managed to achieve successful IT implementation in the case. In trying to understand this problem, the research will seek to ask the following questions:

- What are the general factors affecting IT implementation in the Nigerian public sector?
- What is the situation regarding the use of IT for SCM functions in the Nigerian public sector?
- What are the prevalent human factors affecting IT implementation for SCM functions in the Nigerian public sector?
- How do the human factors interplay to affect IT implementation and SCM functions in the Nigerian public sector?
- How can the human factors be managed to enhance IT implementation and SCM functions in the Nigerian public sector?

3 Choosing from the Broad Research Classifications

Haven understood the objective of the research, there is the need to identify the type of research being proposed from the broad classifications of research methods. There are notably several research classifications owing to the diverse perspectives from which researchers have viewed this concept. Some common categorisations include the distinction between 'quantitative' and 'qualitative', 'objective' versus 'subjective' and also the 'predictive/control' as against 'explanatory/perceptive' approaches. The most widely accepted categorization is the quantitative and qualitative research classifications (Burrell and Morgan, 1979; Luthans and Davis, 1982; Myers, 1997). Quantitative research methods were initially developed for scientists to learn about

natural concepts. It is often argued that quantitative methods are all about numbers and not a common approach for social science research. This is, however, not the case as quantitative methods are widely used for research in the social sciences including studies in IS. The qualitative research methods were however primarily developed to enable social scientists gain knowledge of cultural and social trends. The shift in IS research from the technological to managerial perspectives orchestrated growing interest in the application of qualitative research methods in the field of study (Myers, 2008). This research which focuses on the human (soft) aspects of IT is aligned with this shifting trend and interest for the use of qualitative approaches in IS studies.

Stake (1995) aptly differentiates quantitative and qualitative research from three stand points. First is that quantitative research seeks to 'explain' while qualitative research seeks to 'understand'. Quantitative research takes an impersonal approach whilst qualitative research follows a personal approach. Quantitative research seeks to discover knowledge with the 'why' questions while qualitative research seeks to construct knowledge with the 'what' and 'how' questions (Harling, 2002). This research adopts the personal approach qualitative practice; seeking to understand a phenomenon with a view to constructing knowledge. The approach is appropriate for this research which involves a phenomenon embedded in a complex relation (Harling, 2002). It is applicable to the objective of the study which seeks to construct knowledge in line with the application of 'what' and 'how' questions; built on the assertion of the social construction of reality (Searle, 1995). A major advantage of this research approach is that it creates a close collaboration between the researcher and the participants while allowing the participants to describe their views of reality (Crabtree and Miller, 1999). The researcher is then able to better understand the participants' actions from the vivid description of their views of reality (Lather, 1992; Robottom and Hart, 1993).

4 Philosophical Perspective of the Research

After identifying the general research classification of a research, another key aspect is the need to understand the philosophical perspective of the research. These are the underlying assumptions that constitute what validates research often referred to as the philosophical perspectives or research paradigms. Ontology is considered as the starting point in this aspect from which epistemology and methods are derived from. Ontology simply refers to the specification of conceptualization of a research (Gruber, 1995). Ontological beliefs are concerned with how the phenomenon under investigation is perceived; if the world is thought to be objective or subjective. That is if they are independent of humans or primarily exist through human actions (Orlikowoski and Baroudi, 1991). Ontologically, the interpretive perspective does not view the world as objectively known and unproblematic but as an 'emergent social process' and thus attempts to understand how and why and people

According to Guba and Lincoln (1994) the four paradigms of qualitative research are positivism, post-positivism, constructivism and critical theory. Similarly Orlikowoski and Baroudi (1991) identified three philosophical perspectives of qualitative research based on epistemology. The philosophical assumptions; critical, interpretive and positivist, are considered as guides that influence qualitative research. It, however, needs stating that while these various paradigms and perspectives are significantly, philosophically distinct, it is somewhat difficult in practice to see a clear cut difference between them (Lee, 1989). This study will not essentially follow the positivist's theory which attempts to test theory with the assumption that reality is obvious and can be explained by quantifiable measures (Yin, 2002). Similarly, it does not align with the critical theory which focuses on oppositions, contradictions and conflicts in the modern world (Myers, 1997). This study supports the constructivism perspective (Guba and Lincoln, 1994) which aligns with the interpretivist paradigm (Orlikowoski and Baroudi, 1991); both regarded as sensitising concepts (Blumer, 1954). The concepts merely steer the interest of the researcher towards the direction to look but not necessarily describing what to see (Schwandt, 1998). There are no pre-defined variables; focus is on the entire complexity of man's activity and creativity; reviewing conditions as they unfold (Kaplan and Maxwell, 1994).

This approach tries to comprehend concepts from the perception of people. The researcher is saddled with the need to explain the process of constructing meanings and showing 'what' and 'how' meanings are personified in the language and actions of social actors (Schwandt, 1998). The proponents of this persuasion try to understand the world from the point view of those who live in it (Searle, 1995). The paradigm recognises the importance of the subjective human creation of meaning but does not totally disregard objectivity. The gain of this approach in this study is that it helps to understand issues in the process and context of the system being considered. It facilitates the understanding of the evolving influence of the system on the context as well as the resulting impact on the system by its context (Walsham, 2011).

5 Intricacies of Choosing Research Methods

Similar to the philosophical assumptions regarding qualitative research, there are various qualitative research methods employed in the application of qualitative research. These methods are the basic techniques or strategies employed in line with the philosophical assumptions; for data collection and creation of a research design framework. Four of the common qualitative research methods include grounded theory, action research, ethnography and case study (Myers, 1997).

The proposed method for this study is the case study method. Case study methods provide tools that afford researchers the opportunity to explore and study complex phenomenon within their context and using different data sources (Baxter and Jack, 2008). This allows for the issue to be explored through multiple lenses thereby ensuring the understanding and disclosure of the various components of the subject.

The approach supports the deconstruction and subsequent reconstruction of various phenomena in the study of individuals or organisation. This flexible and rigorous approach facilitates the development of theory, evaluation of programmes and development of interventions (Yin, 2003) akin to the objectives of this research. The use of qualitative case study as proposed by Stake (1995) and Yin (2003, 2006) ensures that the interested phenomenon is well explored and explicit (Baxter and Jack, 2008). Sources of data for qualitative research are often obtained from interview, personal observations, questionnaires etc. These will be employed in this research. There are other factors that influence data collection irrespective of the methods being applied. These include theoretical influences, choice of theory and experiential influence of the researcher. These factors will be enumerated in the subsequent paragraphs in line with the researchers approach in the proposed study.

5.1 Theoretical Influence

There are disagreements as to the role of theory in case study applications. Stake (1995) believes that theory can be absent when focusing on a case and its issues, while Creswell (1994) asserts that theory can be applied at the end of the study in which existing theories can be compared with the study results. Yin (1995), however, underlines that theory can be used to guide the study in a probing manner. This research however agrees with Harling (2002) that existing theories provide direction and structure to a researcher's inquiry. Existing theories will therefore be employed to direct and structure initial inquiry as well as to filter received data to validate the theories. Care would however be taken to prevent existing theories from stereotyping or influencing the research results (Harling, 2002). Consequently, the research will be sensitive to the paradox between the case situations and existing theory.

5.2 Choice of Theory

Walsham (1995) supports Yin's (1995) perspective that theory can be used as an initial guide as part of an iterative process in a study. He also asserts that it can be utilised as a final product in a research. In an analysis of a number of studies undertaken by researchers, Walsham (2005) submits that the choice of theory in a research is essentially subjective. The choices for the theory adopted by the researchers in the cases he reviewed were considered to be predominantly based on their own experiences, background and interests. While bearing in mind the need for the researcher's to choose theory in line with the objective of the research, the key point observed was that the researchers could personally relate with the concepts of the theories they used. On this note the advice for upcoming researchers was for the need to choose theories that the researcher considers insightful; one that really engages the researcher and not necessarily one that is 'fashionable'. It is bearing this in mind that the research ventured into some theories which were apparently not fashionable to some reviewers. The researcher did not dismiss the theories following the receipts of critical comments from reviewers. The comments spurred the researcher to read widely on different theories and also read deeply about the selected theories. Consequently, the basis for the choice of theories reflected those that were inspiring and considered appropriate to gain insight from field data (Walsham, 2005).

5.3 Experiential Influence

There are also diverse views about the application of experiential knowledge in research. The traditional view holds that inputs from a researcher's background and identity are biases that alter a research study. As such, proponents of the view recommend that such inputs should be ignored and eliminated. The classic proponents however assert that experiential data are valuable components of research that should not be ignored. The concept holds that rather than ignore valuable experiential data on the basis that they could bias research results, such data should be critically explored to elicit the potential 'gold' embedded there in (Do et al, 2004). This research appreciates the fears of the traditional proponents in avoiding bias but supports the classic view regarding the benefits harnessing experiential data. The approach will therefore follow the reasoning of Peshkin and Strauss (1988, 1994) that researchers should not suppress their primary experience or be swept away and overwhelmed by it. The strategy rather is to raise its consciousness and use it as part of the inquiry process (Do et al, 2004). The researcher in this case has over 14 years work experience in the proposed organization. The interest in the research evolved haven experienced the problem situation from various perspectives. The understanding of the situation is enhanced by the researcher's involvement as a resource person in the organisation; haven explored the situation through the lenses of academic and field exercises. It is with this experience that the researcher approached literature to gain theoretical knowledge of the subject.

6 Strategies for Understanding Soft Issues of Information Systems

There is evidence in the literature of strategies that have been developed to understand the complexities of the soft/human issues involved in IT development and implementation (Cox, 1999; Goldfinch, 2007; and Walsham, 2011). This section will attempt to identify some critical strategies in this regard in view of their relevance with the objective of this research.

6.1 Effective Technical and Human Implementation of Computer-based Systems

The Effective Technical and Human Implementation of Computer-based Systems (ETHICS) is a socio- technical approached developed by Mumford (1995). Its methodology is based on a participative approach aimed to foster genuine users' participation by providing inputs for the design specification of a project up to the evaluation of the designed prototype (Vidgen and Madsen, 2000). The view holds that technology must fit with social and organisational factors. It propagates improving quality of work life and enhancing users' job satisfaction. This approach sees the

development of computer-based systems from a change management perspective which involves conflicts of interest between the actors involved in the process. It explains that conflicts do not only manifest between management and workers but often exists within the management team as well as amongst the workers. The successful implementation of new systems is therefore a process of negotiation between the affected and interested parties. It incorporates a framework for job satisfaction which includes five dimensions; knowledge fit, psychological fit, efficiency fit, task fit and ethical fit.

6.2 Business Process Re-engineering

Business process re-engineering (BPR) is a management technique promulgated by Hammer (1990) that radically transforms an organisation for dramatic improvement (Akhavan, Jafar and Al-Ahmadi, 2006). Hammer claims that the major challenge of managers was to eliminate all forms of works that do not add value to a process rather than using IT to automate the process. The idea is that operations and processes should be simplified and standardised before automation with IT. Automating existing process with IT may lead to more efficient way of doing wrong things (Hammer and Champy, 1993; Gunasekaran and Nath, 1997). At the heart of BPR is the need for mangers to think differently and break off from old work structures which are considered obsolete to current technological trends. Akhavan, Jafar and Al-Ahmadi (2006) highlight that an estimated 70 per cent of BPR projects fail for varied reasons including internal organisational problems and the view that an IT organisation is an Hammer (1990) attributes BPR failures to poor change obstacle to innovation. management and leadership styles. Roby (1995) clearly wraps up these points in attesting that the management of the human (soft) element is the major aspect of BPR which determines its success or failure. This suggests the need to support BPR with other approaches that will take care of the negative consequences in line with the human aspect considerations.

6.3 Soft Systems Methodology

The Soft Systems Methodology (SSM) is an organised way of tackling ill-defined and 'messy' real world situations. It is a methodology that employs systems thinking which enables it to be highly defined. Nonetheless, it is flexible in use and broad in scope. It uses systems concepts like other systems approaches, such as emergence, control and hierarchy (Checkland and Scholes, 1990). It is distinct in employing the concept of a human activity system as well as in expressing phenomenological models. The approach has already been used extensively by a variety of management practitioners in the UK; including on a good number of computer-based IS projects. It has proved to be exceptionally relevant to IS design and development (Walsham, 2011). It has a very impressive track record and has been successfully used in several different interventions and found to be particularly flexible (Lewis, 1999). SSM which has been undoubtedly influential has also got its fair share of criticisms. It has been criticised for ignoring the constraining effects of existing power relations, and

thus said to be geared to favour the interests of those in authority. Checkland, however, disputes this notion stressing that radical change can be achieved through SSM (Walsham, 2011). SSM is intrinsically a collaborative approach and practical users involve other people in the process of problem handing (Checkland and Scholes, 1990).

7 Research Action Plan

The proposed research employs triangulation which is the use of a combination of more than one investigator, data source, methodological approach, theoretical perspective or data analysis method to investigate a research question (Thurmond, 2001). This approach is considered appropriate in view of the identified strengths and weakness of the identified approaches in the research. For instance, SSM is appropriate for understanding and exploring complex situations (Checkland, 1981). However, studies have shown that the development of conceptual models with SSM is somewhat ambiguous; it does not give a clear relationship between the conceptual model and physical world (Gregory, 1992). BPR on the other hand is quite impressive with reconstructing processes. It is helpful in simplifying and standardising processes; to eliminate work that does not add value to process; thereby developing a model that revamps organisations (Hammer, 1990). BPR however fall short of effectively considering human aspect considerations. On the other hand, ETHICS approach which aims to foster genuine users' participation is valuable in handling human considerations. It is in view of such display of varied strengths and weaknesses that the researcher considers it appropriate to combine theories and methods in their areas of strength.

The research approach will therefore employ techniques from SSM, BPR and ETHICS which are in tandem with the techniques of a collaborative inquiry – action research. The research can however not be considered to be a full blown action research as it does not include the actual implementation of an intervention action (Reason and Bradbury, 2002). It however vividly explores the use of action research techniques bearing in mind the disposition and relationship of the researcher in the case study. The study invariably employs action research techniques which involve a reflective, interactive inquiry process with other stakeholders in the field.

The case study analysis has been broadly categorised into three phases; defining the problem situation, developing a management framework to address the situation and ensuring that the framework is feasible in the context. These three phases were deduced from the 7 stages of SSM. The phases do not imply that the study will be conducted in the sequential order. The identification of these phases primarily helps to facilitate sense making of the inquiry process of the case study. The use of 'CATWOE' and rich pictures (SSM techniques) will be employed to assist defining the problem situation in the first stage of the Nigerian case study. This includes the structured evaluation of those elements which are relatively stable and slow to change. This consists of physical structures like office buildings as well as social

structures like job roles of individuals. The process evaluation includes the prevalent transformation process in the situation with its inputs and outputs (Lewis 1994). The results will be expressed with the SSM technique of drawing rich pictures; pictorially describing the problem situation. The rich picture graphically captures details of structures, processes, climate, people, issues expressed by people as well as conflicts (Williams, 2005). Rich pictures have been proven to be predominantly successful in expressing issues concerning varied parties. They present facts in an astute manner to stimulate discussions to a level that is rarely achievable with only written text or tables of figures (Lewis, 1994).

The next stage of the study which involves the development of the management framework will bear in mind the underpinning values of BPR which emphasizes the need to standardize and simplify processes before automating with IT. The development of the management framework will evolve from the analysis of data obtained from answering the research questions. ETHICS's technique which stresses incorporating user's participation to ensure technology fit with social and organizational factors will be employed at the last stage to guarantee the feasibility and validity of the developed framework in the case.

7 Conclusion

This research corroborates the evidence in the literature of the complex interactions and implications of human factors in IT implementation. There is somewhat limited understanding about the management of the 'softer' and less detectable components of IT and SCM; which have been observed as being difficult to implement. There is thus an apparent gap between the theory and practice of IT and SCM functions; hence, the need for more understanding about its theories and practices. This study looks at the soft (human) aspect considerations of IT implementation in SCM in a Nigerian public sector. The complexities and interrelations of the soft issues necessitate the choice of systems thinking approach. The aim is to undertake a holistic study of the soft issues; managing the problem elements in isolation may cause important interactions to be overlooked. This is important as the solution to a whole scheme is often more than the sum of the solutions of its individual elements. The developed management framework is to enable precise thinking and intervention to enhance IT implementation and SCM functions.

The research hopes to contribute in the areas of policy and practice regarding IT implementation and SCM in the public sector organisation in Nigeria. The contribution to theorization will enable precise thinking and intervention to enhance IS implementation and public sector SCM functions. The results will also contribute to disapproving the myth of case study theorization (Flyvberg, 2006). The combination of methods employed in the research is intrinsically unique. This approach holds promising contributions to knowledge regarding methods and theories for conceptualising and analysing IT implementation and SCM functions to optimise performance and impact.

References

- 1. Akhavan, P., Jafari, M. and Ali-Ahmadi A.R., "Exploring the interdependency between reengineering and information technology by developing a conceptual model", *Business process Management Journal* Vol. 12, No. 4, pp. 517-534 (2006)
- Baxter, P. and Jack, S. "Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers". *The Qualitative Report* Vol. 13 No. 13 pp. 544 – 559 (2008)
- 3. Blumer, H., "What is wrong with social theory": Theory in Marketing, Richard D. Irwin, Inc., Chicago, 1950
- 4. Brooke, C. and Maguire, S. "Systems Development: A Restrictive Practice?" *International Journal of Information Management*, Vol. 18, No. 3, pp. 165–180 (1998)
- Checkland, P. and Scholes, J. "Soft Systems Methodology in Action", John Wiley and Sons Ltd Publications ISBN 0-471-92768-6 (1990)
- Cheikhrouhou, N., Pouly, M., and Huber, C. "An Empirical Study on Human and Information Technology Aspects in Collaborative Enterprise Networks", *Journal of Universal Computer Science*, Vol. 17, No. 2, pp. 203-223 (2011)
- 7. Cox, A., "A research agenda for supply chain and business management thinking Strategy" *Supply Chain Management: An international Journal*, Vol. 4, No. 4, pp.209-211(1999)
- 8. Crabtree, B.F. and Miller, W.L. Second Edition, Doing Qualitative Research, Sage Publications, Thousand Oaks, California (1999)
- 9. Creswell, J. Research Design: Qualitative and Quantitative Approaches, Sage Publications, Thousand Oaks, California (1994)
- 10.Davenport, T. H. Process Innovation–Reengineering Work through Information Technology. Harvard Business School Press, Boston, Massachusetts, 1993
- 11.Do, W., Think, Y. and Going, I, "Conceptual Framework: What do you think is going on?" pp.33–64 (2004)
- 12.Flyvbjerg, B. "Five Misunderstandings About Case-Study Research", Qualitative Inquiry, Vol. 12 No. 2, pp. 219-245, Sage Publications (2006)
- 13.Gichoya, D. "Factors Affecting the Successful Implementation of ICT Projects in Government", *The Electronic Journal of e-Government*, Vol. 3 No. 4, pp.175 184 (2005), available online at www.ejeg.com
- 14.Gregory, WJ. Critical systems thinking and pluralism: a new constellation. City University, London (1992)
- 15. Gruber, T. R., "Toward principles for the design of ontologies used for knowledge sharing", International Journal of Human-Computer Studies, Vol. 43, Issues 4-5, pp. 907-928 (1995)
- 16.Guba, E.G. and Lincoln, Y.S. "Competing paradigms in qualitative research," in Handbook of Qualitative Research, N.K. Denzin and Y.S. Lincoln (eds.), Sage, Thousand Oaks, 1994, pp. 105-117
- 17. Gunaskeran A. and Nath, B. "The role of information technology in business process reengineering" *International Journal of Production Economics* Vol. 50 pp. 91-104 (1997)
- 18.Hammer, M., and Champy, J.: Reengineering the Corporation: A Manifesto for Business Revolution. Harper Business, New York, 1993
- 19.Hammer, H., "Reengineering Work: Don't Automate, Obliterate", Harvard Business Review Vol. 68, No. 4, pp.104–112 (1990)
- 20.Harling, K., "An Overview of Case Study", A paper supporting a discussion conducted at the learning workshop, "Case Studies: Their Future Role in Agricultural and Resource Economics," held at the annual meeting of the American Agricultural Economics Association, Long Beach, California, July 27, 2002
- 21.Kaplan, B. and Maxwell, J.A., "Qualitative Research Methods for Evaluating Computer Information Systems," in Evaluating Health Care Information Systems: Methods and

Applications, J.G. Anderson, C.E. Aydin and S.J. Jay (eds.), Sage, Thousand Oaks, CA, 1994, pp. 45-68 (1994)

- 22.Ko, R.K.L. "A Computer Scientist's Introductory Guide to Business Process Management (BPM)", *Crossroads*, Vol. 15, No. 4, pp.11-18 (2009)
- 23.Lewis, P., Information Systems-Development, Pitman Publications, London, ISBN 0 273 031074 (1994)
- 24.Maguire, S. and Redman, T., "The role of human resource management in information systems development", *Emerald Group Publishing Limited*, Vol. 45 No. 2, pp. 252-264 (2006)
- 25.Mumford, E., "The story of socio-technical design: reflections on its successes, failures and potential", *Info Systems Journal Vol.* 16, pp.317-342 (2006)
- 26.Myers, M. D., "Qualitative Research in Information Systems," *MIS Quarterly* Vol. 21, No. 2, pp. 241-242. MISQ Discovery, archival version, June 1997, http://www.misq.org/discovery/MISQD_isworld/. MISQ Discovery, updated version, last modified: January 4, 2008 <u>http://www.qual.auckland.ac.nz/</u>
- 27. Myers, M. D., "Investigating Information Systems with Ethnographic Research," Communications of the Association for Information Systems Vol. 2, No. 23, pp.1-19 (1999)
- 28.Olorunniwo, F. and Li, X, "Information sharing and collaboration practices in reverse logistics", *Supply Chain Management: An international Journal*, Vol 15 No. 6, pp. 454-462 (2010)
- 29.Orlikowski, W.J and Baroudi, J.J., "Studying information technology in organizations: Research approaches and assumptions", Information Systems Research, Vol. 2, No. 1, pp. 1– 28 (1991)
- 30.Reason, P. and Bradbury, H., Handbook of Action Research. Participative Inquiry and Practice. 1st Edition. London: Sage ISBN 0-7619-6645-5 (2002)
- 31.Schwandt, T.A. "Constructivist, Interpretivist Approaches to Human Inquiry". In: N.K. Denzin and Y.S. Lincoln (eds) *The Landscape of Qualitative Research: Theories and Issues*. London: Sage, pp. 221-259 (1998)
- 32.Stake, R.E. The art of case study research. Thousand Oaks, CA:Sage (1995)
- 33. Thurmond, V.A. "The Point of Triangulation", *Journal of Nursing Scholarship*, Vol.33, No. 3, pp.253-258 (2001)
- 34.Vigen, R., and Madsen, S., "Exploring the Socio-Technical Dimension of Information Systems Development Use Cases and Job Satisfaction", *Socio-Technical approaches to IS Development* (2000)
- 35. Walsham, G., Interpreting Information Systems in Organizations, Wiley, Chichester (2011)
- 36.Walsham, G. and Sahay, S., "Research and Information Systems in Developing Countries: Current Landscape and Future Prospects" Information Technology for Development Vol. 12, No. 1, pp. 7-24 (2006)
- 37.Walsham, G., "Interpretive case studies in IS research: nature and method," European Journal of Information Systems Vol. 4, pp. 74-81(1995)
- 38.Willcocks, L., "Information Technology in Public Sector Settings: Towards Effective Systems", *International Journal of Public Sector Management*, Vol. 2 No. 3, pp.null – null (1993)
- 39.Williams, B., and Foundation, T. K. SOFT SYSTEMS. Methodology, (December), 1-20 (2005)
- 40.Yin, R.K., Case study research: Design and methods (3rd Edition). Thousand Oaks, CA:Sage (2003)

Existing approaches for cross-platform development and deployment of cloud applications

Fotis Gonidis¹, Iraklis Paraskakis¹ and Anthony J. H Simons²

¹ South-East European Research Centre (SEERC),
 City College – International Faculty of the University of Sheffield,
 24 Proxenou Koromila Street, 54622 Thessaloniki, Greece {fgonidis, iparaskakis}@seerc.org

² Department of Computer Science, University of Sheffield, Regent Court, 211 Portobello Street, Sheffield S1 4DP, United Kingdom A.Simons@dcs.shef.ac.uk

Abstract. Cloud computing is a relatively new paradigm that promises to revolutionize the way IT services are provided. However, the issue of cloud application portability could be an impediment for the wide adoption of cloud computing. In this paper, certain existing approaches addressing application portability are presented. Each solution is developed in a different direction. Next we briefly state the scope of our research focus and the future research directions on the field of cross platform development and deployment of cloud applications

Keywords: Cloud Computing, Portability

1 Introduction

The growing interest in cloud application platforms has resulted in a large number of platform offerings being already available on the market. However, different cloud application platform offerings are characterized by considerable heterogeneity. Because of incompatibilities, users that develop applications on a specific platform may encounter significant problems when trying to deploy them in a different environment. This gives rise to the familiar problem of vendor lock-in [1], which has been a challenge long before the advent of cloud computing.

Incompatibilities which may arise while deploying an application across various platforms include the differences in programming languages and databases, specific configuration files and differences in provided services such as the message queue service, file storage service, billing service etc [2]. However, for developers to be able to exploit the full advantages of PaaS [3], they should be able to overcome the conflict points and deploy their cloud applications across multiple platforms, without lock-in to a particular vendor. To achieve this, a new approach to cloud application

development must be adopted. The key concept is for users not to develop applications directly against proprietary platform environments. Rather, developers should use either standard and widely adopted technologies, or abstraction layers which decouple application development from specific target platforms. Ensuring portability across cloud providers would eliminate the vendor lock-in problem and in turn, this would increase consumers' trust towards cloud computing and public cloud services.

This paper examines existing approaches which address the issue of application portability. Rather than providing an exhaustive list of existing approaches we focus on certain approaches which follow a different solution paradigm. Next we briefly state our own research scope and future directions towards addressing the issue of application portability.

2. Existing approaches

Significant work has been done in order to minimize the vendor lock-in effect and allow the development of applications which can be deployed across multiple platforms. In this section we briefly state some prominent approaches that adopt different solution architecture.

1) Wrapper approach. The wrapper approach consists of an API library which abstracts the access to the native APIs offered by the specific cloud service providers. jClouds [4] is a major example. It provides abstraction for major cloud storage services such as Amazon S3 and Microsoft Azureblob. It allows the application to perform actions on files located in remote stores which are offered by a cloud provider.

From a technical point of view, jClouds is a client implementation of the REST API [5] which cloud providers expose to allow access to their storage service. Therefore developers can use a single method offered by the library to access any of the supported storage services.

Alternatively, developers should implement their own HTTP requests for each storage service or use the client implementations offered by the cloud providers. In both cases a significant development effort is needed when deploying the application across different cloud platforms. jClouds reduces the effort by introducing an abstraction layer between the application and the storage service. Similar solution which offers a wrapper API library, is the Simple Cloud API [6].

2) Adoption of common standards. Instead of using a wrapper library to abstract the access to the native cloud service APIs, users can choose cloud platforms that adopt common standards and thus avoid any lock-in. Openshift is such a cloud platform which does not use any proprietary API [7]. Another example of an open source platform contributing to application portability is the European project mOSAIC [8]. mOSAIC adopts standard and widely used technologies. On top of that the platform offers the so called interoperability API [9]. Developers can exploit the interoperability API in order to access resources such as databases, file storage and message queuing service from various cloud providers. However, the range of the supported resources is significantly limited.

While the cloud platforms which adopt common standards minimise the lock-in effect, they usually do not offer a range of services that users could exploit in order to enrich their applications. mOSAIC allows a certain access to cloud resources such as file storage, databases, messaging system, but the developers are required to use a certain programming concept which increases the learning curve of the platform.

3) Model-driven engineering (MDE) approach. Another approach towards achieving cloud application portability is to exploit MDE techniques. Using MDE, the application creation process begins with building a platform independent model (PIM). Then, using automated model transformations, the PIM is translated into a platform specific model (PSM) targeting particular cloud platforms.

MobiCloud [10] is an example which leverages the benefits of model-driven engineering (MDE) field. MobiCloud is a domain specific language (DSL) that allows developers to create simple Create, Read, Update, Delete (CRUD) applications using a graphical editor and a scripting language. The platform automatically generates the source code and the required configuration files for uploading the application on Google App Engine and Amazon EC2.

Although MobiCloud enables the development and deployment of cross platform applications, it has a limited application scope. Developers can only create simple CRUD applications and target a limited set of cloud platforms.

Additional examples which leverage MDE techniques are [11] [12], which both propose the definition of a modeling language for describing cloud applications.

3. Future research directions

In the previous section we mentioned certain existing approaches, each one of them addressing the issue of application portability from a different perspective. The first one offers a third party library which abstracts certain storage service providers. The second one constitutes an open source cloud platform where applications can be deployed and executed. The third one leverages model driven engineering techniques and offers an online development environment for creating simple CRUD applications. Each one of the approaches involves certain advantages and disadvantages as mentioned in section 2. In this section we briefly state our own future research directions in the field of cross-platform development and deployment of cloud applications.

Our intention is to create an abstraction framework which can be used by developers to allow access to specific cloud services independent of the cloud provider. The target cloud services can be a message queue service or a platform specific service such as the billing service. Our motivation for choosing this scope is that, to the best of our knowledge, this is a relatively unexplored area and there is no or little related work on the field. In addition to that, the functionality of the framework can be enriched by leveraging existing solutions, such as the jClouds storage service.

An initial high-level overview of the envisioned solution is shown in Figure 1. The abstraction framework will consist of two parts. The first one will contain platform

independent models which will represent the cloud services that the framework will support. For each supported service a set of models will be available. A domain specific language can be designed to leverage the use of the models. The second part will consist of concrete adapters which will be responsible for transforming the abstract models into concrete source code for the target cloud platform such as Google App Engine, Amazon Elastic Beanstalk etc.

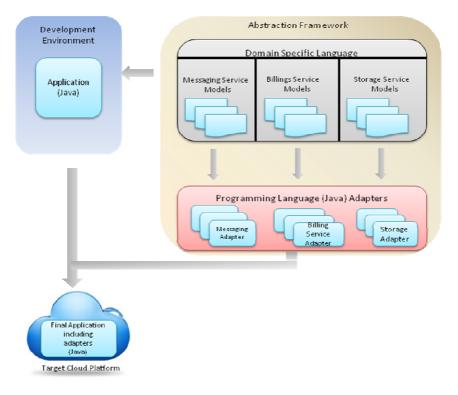


Fig. 1. High level architecture of the abstraction framework

The developer will initiate the development of an application using a popular development environment such as Eclipse and a programming language such as Java. When the application needs to use a cloud service which is supported by the framework, the developer will use the domain specific language in order to configure the service and model the desired functionality. Then, the source code for the specific cloud platform will be generated, using the appropriate adapter. The generated source code will be merged with the rest of the application and deployed on the target platform.

The contribution of the proposed research directions will be threefold. First, it will systematically explore the challenges involved when trying to abstract cloud service providers. Second, it will attempt to propose a formal methodology and best practices for abstracting cloud services. Third, the envisioned abstraction framework will attempt to expand its functionality by leveraging existing solutions. For example, a storage service can be offered by integrating the jClouds framework.

4. Conclusions

In this paper the issue of cross platform development and deployment of cloud applications was introduced. Certain existing approaches, addressing the issue of application portability, were presented. Each approach follows a different solution direction ranging from a wrapper library, to adopting common standards and MDE techniques. Next, our research directions towards addressing application portability were briefly stated. Our intention is to focus on areas that have not been yet extensively explored, such as cloud messaging services or billing services. Additionally, we intend to explore the benefits of introducing a domain specific language for allowing easy access and configuration of the offered services.

References

- 1. Neal Leavitt, "Is Cloud Computing Really Ready for Prime Time?", *Computer*, vol. 42, no. 1, pp. 15–20, Jan. 2009.
- 2. Fotis Gonidis, Iraklis Paraskakis, and Dimitrios Kourtesis, "Cloud Application Portability. An Initial View", in *6th Balkan Conference in Informatics*, Thessaloniki, Greece, 2013.
- 3. F. Gonidis, I. Paraskakis, and D. Kourtesis, "Addressing the Challenge of Application Portability in Cloud Platforms", in *7th South-East European Doctoral Student Conference*, Thessaloniki, 2012, pp. 565–576.
- "jclouds(May 2013)", [Online], Available: http://jclouds.incubator.apache.org/, [Accessed: 10-May-2013].
- R. T. Fielding and R. N. Taylor, "Principled design of the modern Web architecture", in Proceedings of the 22nd international conference on Software engineering, New York, NY, USA, 2000, pp. 407–416.
- "SimpleCloud(May 2013", [Online], Available: http://www.simplecloud.org/, [Accessed: 10-May-2013].
- "Openshift(Jun 2013)", [Online], Available: https://www.openshift.com/, [accessed: 30-Jun-2013].
- 8. D. Petcu, C. Craciun, M. Neagul, I. Lazcanotegui, and M. Rak, "Building an interoperability API for Sky computing", in 2011 International Conference on High Performance Computing and Simulation (HPCS), 2011, pp. 405–411.
- 9. Gorka Echevarria, "Description of mOSAIC's API", D1.3, Aug. 2011.
- 10. A. Ranabahu and A. Sheth, "Semantics Centric Solutions for Application and Data Portability in Cloud Computing", in *Cloud Computing Technology and Science (CloudCom)*, 2010 IEEE Second International Conference on, 2010, pp. 234–241.
- 11. J. Esparza-Peidro and F.D. Munoz-EscoQ, "Towards the Next Generation of Model-Driven Cloud Platforms", TR-ITI-SIDI-2011/001, Valencia, 2011
- M. Hamdaqa, T. Livogiannis, and L. Tahvildari, "A reference model for developing cloud applications," in Proceedings of the 1st International Conference on Cloud Computing and Services Science, 2011, pp 98-103

High-Throughput Asymmetric Encryption cryptanalysis in Embedded Systems

Marsida Ibro¹, Marin Aranitasi², Blerta Moçka³

¹Faculty of Information Technology, Polytechnic University of Tirana ²Center for R&D on Faculty of Information Technology, Polytechnic University of Tirana, ³Faculty of Natural Sciences, University of Tirana marsida_ibro@hotmail.com, m_aranitasi@hotmail.com, mocka.blerta@gmail.com

Abstract. Most of the advanced security systems rely on public-key schemes based either on the factorization or the discrete logarithm problems. The both problems are closely related, a major breakthrough in cryptanalysis tackling one of those problems could render a large set of cryptosystems completely useless. The McEliece scheme is based on the alternative security assumption that decoding unknown linear binary codes is NP-complete. In this work, we investigate the efficient implementation of the McEliece scheme on embedded systems what was – up to date – considered scheme on embedded systems a challenge due to the required storage of its large keys. In this work we have explained the McEliece scheme implementation on a low-cost 8-bit AVR microprocessor suitable for many embedded system applications.

Keywords: Cryptography, Public Key, Embedded systems, AVR microprocessor

1 Introduction

The advanced properties of public-key cryptosystems are required for many cryptographic services, such as key establishment between parties and digital signatures. RSA, ElGamal, and later ECC have evolved as most popular choices and build the foundation for virtually all practical security protocols and implementations with requirements for public-key cryptography.

Alternative schemes namely the Merkle signature scheme and cryptosystems based on multivariate quadratics, provide digital signatures but, unlike RSA and ECC, do not provide public key encryption. In this work, we demonstrate the efficient implementation of another public-key cryptosystem proposed by Robert J. McEliece in 1978. Unlike RSA, ElGamal, and ECC that rely on two primitive security assumptions, namely the factoring problem (FP) and the discrete logarithm problem (DLP), the McEliece cryptosystem is based on coding theory [1]. The McEliece cryptosystem incorporates a linear error-correcting code (namely a Goppa code) which is hidden as a general linear code. For Goppa codes, fast decoding algorithms exist when the code is known, but decoding codewords without knowledge of the coding scheme is proven NP-complete [2]. Accordingly the McEliece scheme is,

contrary to DLP and FP-based systems, also suitable for post-quantum era since it will remain unbroken when appropriately chosen security parameters are used [3]. McEliece makes a good addition to the signature schemes, as it provides asymmetric encryption. Fundamental operations for McEliece are based on encoding and decoding binary linear codes in binary extension fields that, in particular, can be implemented very efficiently in dedicated hardware. Unlike FP and DLP-based cryptosystems, operations on binary codes do not require computationally expensive multi-precision integer arithmetic what is beneficial for small computing platforms [4].

Algorithm 2 McEliece Message Encryption

Require: m, $K_{pub} = (\bar{G}, t)$ Ensure: Ciphertext c1: Encode the message m as a binary string of length k2: $c` \leftarrow m \cdot \bar{G}$ 3: Generate a random n-bit error vector z containing at most t ones 4: c = c` + z5: **return** c

Algorithm 3 McEliece Message Decryption

Require: c, $K_{sec} = (P^{-1}, G, S^{-1})$ Ensure: Plaintext m 1: Encode the message m as a binary string of length k to decode c` to m`= m·S 2: c` $\leftarrow c \cdot P^{-1}$ 3: Use a decoding algorithm for the code C 4: $m \leftarrow m` \cdot S^{-1}$ 5: **return** m

Note that Algorithm 2 only consists of a simple matrix multiplication with the input message and then distributes t random errors on the resulting code word. Thus, the generation of random error vectors requires an appropriate random number generator to be available on the target platform. The most recent implementation of the McEliece scheme is due to Biswas and Sendrier and presented a slightly modified version for PCs that achieves about 83 bit security (taken the attack in [3] into account). Comparing their implementation to other public key schemes, it turns out that McEliece encryption is even faster than that of RSA and NTRU [5].

The McEliece scheme is a public key cryptosystem based on linear error-correcting codes. The secret key is the generator matrix *G* of an error-correcting code with dimension *k*, length *n* and error correcting capability *t*. To create a public key, McEliece defined a random $k \times k$ -dimensional scrambling matrix *S* and $n \times n$ -dimensional permutation matrix *P* disguising the structure of the code by computing $\bar{G} = S \times G \times P$. Using the product key $K_{pub} = (\bar{G}, t)$ and private $K_{sec} = (P^{-1}, G, S^{-1})$, encryption and decryption algorithms can be given by Algorithm 2 and Algorithm 3, respectively.

As mentioned in the introduction, the main caveat against the McEliece cryptosystem is the significant size of the public and private key. The choice of even a minimal set of security parameters (m = 10, n = 1024, t = 38, k \ge 644) according to [6] already translates to a size of 80.5 kB for the public key and at least 53 kB for the private key (without any optimizations). However, this setup only provides the comparable security of a 60 bit symmetric cipher. For appropriate 80 bit security, even larger keys, for example the parameters m = 11, n = 2048, t = 27, k \ge 1751, are required.

1.1 Goppa Code

Theorem 1. Let G(z) be an irreducible polynomial of degree t over F_{2m} . Then the set

$$\Gamma(G(z), \mathbb{F}_{2^m}) = \left\{ (c_\alpha)_{\alpha \in \mathbb{F}_{2^m}} \in \{0, 1\}^n \mid \sum_{\alpha \in \mathbb{F}_{2^m}} \frac{c_\alpha}{z - \alpha} \equiv 0 \right\}$$
(1)

defines a binary Goppa code C of length $n = 2^m$, dimension $k \ge n - mt$ and minimum distance $d \ge 2t + 1$. The set of the c_{α} is called the support of the code. A fast decoding algorithm exists with a runtime of n t.

For each irreducible polynomial G(z) over F_{2m} of degree t exists a binary Goppa code of length n = 2m and dimension k = n - mt. This code is capable of correcting up to t errors [7] and can be described as a $k \times n$ generator matrix G such that

 $C = \{mG: m \in F_2^k\}$. To encode a message m into a codeword c, represent the message m as a binary string of length k and multiply it with the k × n matrix G.

However, decoding such a codeword r on the receiver's sidewith a (possibly) additive error vector e is far more complex. For decoding, we use Patterson's algorithm [8] with improvements from [9].

Since $\underline{r} = \underline{c} + \underline{e} \equiv \underline{e} \mod G(z)$ holds, the syndrome Syn(z) of a received codeword can be obtained from equation (1) by:

$$Syn(z) = \sum_{\alpha \in \mathbb{F}_{2m}} \frac{r_{\alpha}}{z - \alpha} \equiv \sum_{\alpha \in \mathbb{F}_{2m}} \frac{e_{\alpha}}{z - \alpha} \mod G(z)$$
(2)

finally recover need to solve the key equation То e, we $\sigma(z) \cdot Syn(z) \equiv \omega(z) \mod G(z)$, where $\sigma(z)$ denotes a corresponding error-locator polynomial and $\omega(z)$ denotes an error-weight polynomial. Note that it can be shown that $\omega(z) = \sigma(z)'$ is the formal derivative of the error-locator and by splitting $\sigma(z)$ into even and odd polynomial parts $\sigma(z) = \alpha(z)^2 + z \cdot b(z)^2$, we finally determine the following equation which needs to be solved to determine error positions:

$$Syn(z)[a(z)^{2} + z \cdot b(z)^{2}] \equiv b(z)^{2} \mod G(z)$$
(3)

To solve Equation (3) for a given codeword r, the following steps have to be performed:

i. From the received codeword r compute the syndrome Syn(z) according to

Equation (2). This can also be done using simple table-lookups.

- ii. Compute an inverse polynomial T(z) with $T(z) \cdot Syn(z) \equiv 1 \mod G(z)$ (or provide a corresponding table). It follows that $[T(z) + z]b(z)^2 \equiv a(z)^2 \mod G(z)$.
- iii. There is a simple case if $T(z) = z \Rightarrow a(z) = 0, b(z)^2 \equiv z \cdot b(z)^2 \cdot Syn(z) \mod G(z) \Rightarrow 1 \equiv z \cdot Syn(z) \mod G(z)$ what directly leads to $\sigma(z) = z$. Contrary, if $T(z) \neq z$, compute a square root R(z) for the given polynomial $R(z)^2 \equiv T(z) + z \mod G(z)$. Based on a observation by Huber [10] we can then determine solutions a(z), b(z) satisfying

$$a(z) = b(z) \cdot R(z) \mod G(z) \tag{4}$$

Finally, we use the identified a(z), b(z) to construct the error-locator polynomial $\sigma(z) = a(z)^2 + z \cdot b(z)^2$

iv. The roots of $\sigma(z)$ denote the positions of error bits. If $\sigma(\alpha i) \equiv 0 \mod G(z)$ with α_i being the corresponding bit of a generator in $\mathbb{F}_{2^{11}}$, there was an error in the position i in the received codeword that can be corrected by bit-flipping.

This decoding process, as required in Step 2 of Algorithm 3 for message decryption, is finally summarized in Algorithm 4.

Algorithm 4 Decoding Goppa Codes

Require: Received codeword *r* with up to *t* errors, inverse generator matrix *iG* Ensure: Recovered message m` 1: Compute syndrome Syn(z) for codeword *r* 2: $T(z) \leftarrow Syn(z)^{-1} \mod G(z)$ 3: **if** T(z) = z **then** 4: $s(z) \leftarrow z$ 5: **else** 6: $R(z) \leftarrow \sqrt{T(z) + z}$ 7: Compute a(z) and b(z) with $a(z) \equiv b(z) \cdot R(z) \mod G(z)$ 8: $\sigma(z) \leftarrow a(z)^2 + z \cdot b(z)^2$ 9: **end if** 10: Determine roots of s(z) and correct errors in *r* which results in *r*` 11: m` \leftarrow r` · iG {Map r_{cor} to m`} 12: return m`

All security parameters for cryptosystems are chosen in a way to provide sufficient protection against the best known attack. A recent paper [3] by Bernstein et.al presents a state-of-the-art attack of McEliece making use of a list decoding algorithm [11] for binary Goppa codes. This attack reduces the binary work factor to break the original McEliece scheme with a (1024, 524) Goppa code and t = 50 to $2^{60.55}$ bit operations. According to [3], Table 1 summarizes the security parameters for specific security levels.

Security Level	Parameters (n, k, t), errors added	Size <i>K_{pub}</i> in Kbits	Size K_{sec} (G(z), P, S) in Kbits
Short-term (60 bit)	(1024, 644, 38), 38	644	(0.38, 10, 405)
Mid-term (80 bit)	(2048, 1751, 27), 27	3, 502	(0.30, 22, 2994)
Long-term (256 bit)	(6624, 5129, 115), 117	33, 178	(1.47, 104, 25690)

Table 1: Security of McEliece depending on parameters.

2 Design criteria for embedded systems

Mid-term security parameters for public key crypto systems providing a comparable security to 64-80 key bits of symmetric ciphers are often regarded. Hence, the implementations are designed for security parameters that correspond to an 80 bit key size of a symmetric cipher. A second important design requirement is the processing and storage of the private key solely on-chip so that all secrets are optimally never used outside the device. With appropriate countermeasures to prevent data extraction from on-chip memories, an attacker can then recover the private key only by sophisticated invasive attacks. For this purpose, AVR μ Cs provide a lock-bit feature to enable write and read/write protection of the Flash memory [12].

Note that larger security parameters of the McEliece scheme are still likely to conflict with this requirement due to the limited amount of permanent on-chip memories of today's embedded platforms.

Analyzing McEliece encryption and decryption algorithms, the following arithmetic components are required supporting computations in F_{2m} : a multiplier, a squaring

unit, calculation of square roots, and an inverter. Furthermore, a binary matrix multiplier for encryption and a permutation element for step 2 in Algorithm 2 are needed. Many arithmetic operations in McEliece can be replaced by table lookups to significantly accelerate computations at the cost of additional memory. For both implementations in this work, our primary goal is area and memory efficiency to fit the large keys and required lookup-tables into the limited on-chip memories of our embedded target platforms.

To make McEliece-based cryptosystems more practical we have to reduce the key sizes. Using a naïve approach in which the support of the code is the set of all elements in \mathbb{F}_{2m} in lexicographical order and both matrices *S*, *P* are totally random, the public key $G^{:}=S \cdot G \cdot P$ becomes a random $n \times k$ matrix. However, since *P* is a sparse permutation matrix with only a single 1 in each row and column, it is more efficient to store only the positions of the ones, resulting in an array with $n \cdot m$ bits.

Another trick to reduce the public key size is to convert G' to systematic form $\{I_k | Q\}$, where I_k is the $k \times k$ identity matrix. Then, only the $(k \times (n - k))$ matrix Q is published [13].

In the last step of code decoding (Algorithm 4), the k message bits out of the n (corrected) ciphertext bits need to be extracted. Usually, this is done by a mapping matrix iG with $G \cdot iG = I_k$. But if G is in systematic form, then this step can be omitted, since the first k bits of the corrected ciphertext corresponds to the message bits. Unfortunately, G and G' cannot both be systematic at the same time, since then

 $G' = \{I_k \mid Q\} = S \times \{I_k \mid Q\} \times P$ and S would be the identity matrix which is inappropriate for use as the secret key.

For reduction of the secret key size, we chose to generate the large scrambling matrix S on the-fly using a cryptographic pseudo random number generator (CPRNG) and a seed. During key generation, it must be ensured that the seed does not generate a singular matrix S. Depending on the target platform and available cryptographic accelerators, there are different options to implement such a CPRNG on embedded platforms. However, the secrecy of S⁻¹ is not required for hiding the secret polynomial G(z) [13].

3 Implementation on AVR microprocessors

In this section, we discuss our implementation of the McEliece cryptosystem for 8-bit AVR microprocessors, a popular family of 8-bit RISC microprocessors (μ C) used in embedded systems. The Atmel AVR processors operate at clock frequencies of up to 32 MHz, provide few kBytes of SRAM, up to hundreds of kBytes of Flash program memory, and additional EEPROM or mask ROM. For our design, we chose an ATxMega192A1 μ C due to its 16 kBytes of SRAM and the integrated crypto accelerator engine for DES and AES [12]. The crypto accelerator is particularly useful for a fast implementation of a CPRNG that generates the scrambling matrix S⁻¹

on-the-fly. Arithmetic operations in the underlying field $\mathbb{F}_{2^{11}}$ can be performed efficiently with a combination of polynomial and exponential representation. We store the coefficients of a value a \mathbb{I} $\mathbb{F}_{2^{11}}$ in memory using a polynomial basis with natural order. Given an $a = a_{10}\alpha^{10} + a_9 \alpha^9 + a_8 \alpha^8 + \ldots + a_0\alpha^0$, the coefficient a is determined by bit i of an unsigned 16 bit integer where bit 0 denotes the least significant bit. In this representation, addition is fast just by performing an exclusive-or operation on 2×2 registers. For more complex operations, such as multiplication, squaring, inversion and root extraction, an exponential representation is more suitable. Since every element except zero in $\mathbb{F}_{2^{11}}$ can be written as a power of some primitive element α , all elements in the finite field can also be represented by α^i with i \mathbb{J} Z_{2m-1} .

Multiplication and squaring can then be performed by adding the exponents of the factors over Z_{2m-1} such as:

$$c = a \cdot b = \alpha^{i} \cdot \alpha^{j} = \alpha^{i+j} | a, b \in \mathbb{F}_{2^{11}}, 0 \le i, j \le 2^{m} - 2$$
(5)

If one of the elements equals zero, obviously the result is zero. The inverse of a value $d\mathbb{I} \mathbb{F}_{2^{11}}$ in exponential representation $d = \alpha^i$ can be obtained from a single subtraction in the exponent $d^{-1} = \alpha^{2^{11-1-i}}$ with a subsequent table-lookup. Root extraction, i.e., given a value $\alpha = \alpha^i$ to determine $r = \alpha^{i/2}$ is simple, when i is even and can be performed by a simple right shift on index i. For odd values of i, m-1 = 10 left shifts followed by a reduction with 2^{11} -1 determine the square root. To allow for efficient conversion between the two representations, we employ two pre computed tables that enable fast conversion between polynomial and exponential representation. Each table consists of

2048 11-bit values that are stored as a pair of two bytes in the program memory. Hence, each lookup table consumes 4 kBytes of Flash memory. Accessing the table directly from Flash memory significantly reduces performance, but allows migration to a cheaper device with only 4 kBytes of SRAM. For multiplication, squaring, inversion, and root extraction, the operands are transformed on-the-fly to exponential representation and reverted to the polynomial basis after finishing the operation.

Due to the large demand for memory, we need to take care of some peculiarities in the memory management of the AVR microprocessor. Since originally AVR microprocessors supported only a small amount of internal memory, the AVR uses 16 bit pointers to access its Flash memory. Additionally, each Flash cell comprises 16 bit of data, but the μ C itself can only handle 8 bit. Hence, one bit of this address pointer must be reserved to select the corresponding byte in the retrieved word, reducing the maximal address range to 64 KByte (or 32K 16 bit words). To address memory segments beyond 64K, additional RAMP-registers need to be used. Additionally the used avr-gcc compiler internally treats pointers as signed 16 bit integer halving again the addressable memory space. For this reason, all arrays larger than 32 Kbyte need to be split into multiple parts resulting in an additional overhead in the program code.

As in the case of the AVR implementation, we decided to implement less intensive operations of the field arithmetic using precomputed log- and antilog tables which are stored in dedicated memory components (BRAM) of the FPGA. Because of these pre computed tables, the number of computational units in hardware can be reduced, also reducing the number of required Lookup-Tables (LUT) in the Configurable Logic Blocks (CLB) of the FPGA. However, note that only 32 BRAMs are available on the Spartan-3AN 1400 FPGA, preventing to have more than one instance of each table for enabling parallel access. Hence, lookups to these tables need to be serialized in most cases. Since the runtime of polynomial multiplication and polynomial squaring is crucial for the overall system performance (Steps 7 and 8 of Algorithm 4), we opted for a parallel polynomial multiplier instead of using the log and antilog tables as well. The remaining, time-critical component is the polynomial inverter which is used in step 1 and step 2 of Algorithm 4, for example to compute the parity check matrix H on-the-fly.

	Resource	Encryption	Decryption	Available
μC	SRAM Flash Memory External Memory	512 Byte 684 Byte 438 Kbyte	12 Kbyte 130.4 Kbyte	16 Kbyte 192 Kbyte
FPGA	Slices LUTs FFs BRAMs Flash Memory	668 (6%) 1044 (5%) 804 (4%) 3 (9%) 4,644 Kbits	11,218 (100%) 22,034(98%) 8,977 (40%) 20 (63%) 4,644 Kbits	11,264 22,528 22,528 32 16,896 Kbits

Table 2: Implementation results of the McEliece scheme with $n = 2048$, $k =$	1751,
$t = 27$ on the AVR ATxMega192 μ C and Spartan-3AN XC3S1400AN-5 FPGA	

4 Conclusions

We now present the results for our two McEliece implementations providing 80 bit security (n = 2048, k = 1751, t = 27) for the AVR 8-bit microprocessor and the Xilinx Spartan-3AN FPGA. We report performance for the ATxMega192A1 obtained from the avr-gcc compiler v4.3.2 and a Xilinx Spartan-3AN FPGA. The resource requirements for our μ C design and FPGA implementation after place-and-route (PAR) are shown in Table 2.

In FPGA design, the CPRNG to generate S⁻¹ based on the PRESENT block cipher turns out to be a bottleneck of our implementation since the matrix generation does not meet the performance of the matrix multiplication. By replacing the CPRNG by a more efficient solution, we can untap the full performance of our implementation. The public-key cryptosystems RSA-1024 and ECC-P160 are assumed 1to roughly achieve a similar margin of 80 bit symmetric security. Note that for ECC are obtained from the ECDSA signature scheme. Embedded implementations of other alternative public key encryption schemes are very rare. The proprietary encryption scheme NTRUEncrypt has received some attention. An embedded software implementation of the related NTRUSign performs one signature on an ATMega128L clocked at 7,37 MHz in 619 ms. However, comparable performance figures of NTRU encryption and decryption for the AVR platform are not available. Note that all throughput figures are based on the number of plaintext bits processed by each system and do not take any message expansion in the ciphertext into account.

References

- 1. R. J. McEliece. A Public-Key Cryptosystem Based On Algebraic Coding Theory.
- Deep Space Network Progress Report, 44:114–116, January 1978.E. Berlekamp, R. McEliece, and H. van Tilborg. On the inherent intractability of certain coding problems. IEEE Transactions on Information Theory, 24(3):384–386, May 1978.
- Daniel J. Bernstein, Tanja Lange, and Christiane Peters. Attacking and Defending the McEliece Cryptosystem. In Proceedings of the International Workshop on PostQuantum Cryptography – PQCrypto '08, volume 5299 of LNCS, pages 31–46, Berlin, Heidelberg, 2008. Springer-Verlag.
- Thomas Eisenbarth, Tim Güneysu, Stefan Heyse, and Christof Paar. MicroEliece: McEliece for Embedded Systems. In Workshop on Cryptographic Hardware and Embedded Systems – CHES 2009, 2009. to appear.
- 5. Daniel J. Bernstein and Tanja Lange. eBACS: ECRYPT Benchmarking of Cryptographic Systems.
- 6. A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. Handbook of AppliedCryptography. CRC Press, New York, 1996.
- E. R. Berlekamp. Goppa codes. IEEE Trans. Information Theory, IT-19(3):590–592, 1973.

- 8. N. Patterson. The algebraic decoding of Goppa codes. Information Theory, IEEE Transactions on, 21:203–207, 1975.
- Y. Sugiyama, M. Kasahara, S. Hirasawa, and T. Namekawa. A Method for Solving Key Equation for Decoding Goppa Codes. Information and Control, IEEE Transactions on, 27:87–99, 1975.
- K. Huber. Note on decoding binary Goppa codes. Electronics Letters, 32:102– 103, 1996.
- 11. Daniel J. Bernstein. List Decoding for Binary Codes. Technical report, 2008
- 12. Atmel Corp. 8-bit XMEGA A Microprocessor. User Guide
- 13. D. Engelbert, R. Overbeck, and A. Schmidt. A summary of McEliece-type cryptosystems and their security, 2007.

Semantic Web Trends on Reasoning Over Sensor Data

Edmond Jajaga¹, Lule Ahmedi² and Lejla Abazi Bexheti¹

¹ South East European University, Faculty of Contemporary Sciences and Technologies, Ilindenska n. 335, 1200 Tetova, Macedonia {e.jajaga, l.abazi}@seeu.edu.mk
² University of Prishtina, Department of Computer Engineering, Nene Tereze nn, 10000 Prishtine, Kosova lule.ahmedi@uni-pr.edu

Abstract. One of the most challenging features of Semantic Web applications is the reasoning module. Based on the Semantic Web paradigm, the tendency of its community is to build applications with well known standards or recommendations. OWL and SWRL are the first to be considered while expressing data semantics. Their support of monotonic reasoning and open world assumption is not always fruitful on the domain of sensor data. Sensor data, which are the case of this study, are specific in terms of their dynamic nature. Addressing the reasoning issues over the gigantic flow of sensor data the community has considered different approaches mostly resulting by building hybrid systems. This paper outlines preliminary work on reasoning issues over sensor data while describing the approaches that have been taken so far focusing on the discipline of water quality management.

Keywords: sensor data, Semantic Web, reasoning, rules, OWL, SWRL, Semantic Sensor Web

1 Introduction

Semantic Web applications are growing day to day. Meanwhile Semantic Web standards are also maturing. Sensor rapid development and deployment in different disciplines including weather forecasting, water quality management, civic planning for traffic management etc. requires efficient machine communication. Many organizations and institutions have taken initiatives to take advantage from the synthesis of both "worlds" to provide semantics on different application domains. In 2008, Kno.e.sis¹ initiated a project for building Semantic Sensor Web assembling sensor metadata from all over the world. The initiative is aligned-well with standardization efforts of W3C and Open Geospatial Consortium (OGC), in particular

¹ The Ohio Center of Excellence in Knowledge-enabled Computing at the Wright State University, http://knoesis.wright.edu/

with Semantic Web² and Sensor Web Enablement³ (SWE) activities, respectively. In fact, Semantic Sensor Web represents a synergy of both initiatives by semantic annotating of simple sensor data i.e. time, spatial and thematic data. In line with Semantic Sensor Web the W3C Semantic Sensor Network Incubator group (the SSN-XG) recently produced an OWL 2 [11] ontology named SSN⁴ [4], which enhances OGC SWE simple spatial and temporal concepts with semantic annotation for analyzing and Linked Data publishing. The SSN ontology models sensor data in four main perspectives: sensor, observation, system and feature and property perspectives.

Sensor data are an example of stream data which are rapidly changing data. These huge amounts of data need to be quickly consumed and reasoned over. For example, if a particular water quality parameter drops from its allowed threshold then this information needs to be captured quickly and an appropriate decision should follow. Sensors continually produce water quality parameter values. Historical and real-time data produced by sensors require a flexible knowledge management system. An area which deals with continuous execution of queries over stream data is Data Stream Management Systems (DSMS). As indicated in [22] it lacks the ability to reason about complex tasks and lacks a protocol for wide publication. The Semantic Web fulfills these gaps but caching all the knowledge for rapidly changing information is inappropriate. Similar to DSMS is Complex Event Processing (CEP) which provides on-the-fly analysis of event streams, but cannot perform reasoning tasks [15]. Following the pros and cons of DSMS and CEP a new research area has been investigated by the community, namely Stream Reasoning [22]. Stream Reasoning integrates data streams, the Semantic Web and reasoning techniques into a unique platform. Unlike in a traditional reasoning environment, where all the information is taken into account, in stream reasoning there are two concepts which indicate the distinguished approach. The window concept restricts the reasoning to a certain subset of statements recently observed on the stream while previous information is ignored, furthermore continuous processing means continuous evaluation of streams against the knowledge base which is constantly changing.

In general, querying RDF triples of stream data has been leveraged with different SPARQL extensions like: Streaming SPARQL [24], Continuous SPARQL (C-SPARQL) [23] and Time-Annotated SPARQL [25].

This study is focused on the Semantic Web rule layer. State-of-the-art rule-based systems for dealing with sensor data reasoning are mainly:

- *Hybrid systems* e.g. CEP with Semantic Web in [14] and [15], production rules with Semantic Web in [13] and [27].
- *Pure Semantic Web rule systems* as given in [5], [6] and [21], but which do not deal with the streaming nature of sensor data.

This paper is organized as follows. Section 2 presents the current trends on ontology and query processing for sensor networks. Section 3 describes the current state-of-the-art of rule-based implementations with focus on sensor application areas.

² See http://www.w3.org/2001/sw/

³ See http://www.opengeospatial.org/domain/swe

⁴ Available at http://purl.oclc.org/NET/ssnx/ssn

The main discussion takes place on Section 4. Finally, the paper is concluded with a discussions section.

2 Ontologies and Queries

Ontologies are defined as formal specifications of a shared conceptualization [19]. Because of its knowledge reuse and sharing, the ontological knowledge model has been widely leveraged for representing wireless sensor networks (WSNs). One of the first WSNs which has seen benefits from including the ontological model into its knowledge base is OntoWEDDS [9], a decision-support system for wastewater management, which extends its previous version's case-based and rule-based reasoning with the WaWO [10] ontology. The evaluation results have yielded an improvement of 70-100% successful diagnosis and no impasse situations including WaWO reasoning, against 60-70% and 10 out of 57 impasse situations without using it.

Interoperability between sensors and sensing systems was enabled with the development of the SSN ontology. Its foundation is based on the DOLCE-UltraLight⁵ (DUL) ontology. To model a knowledge base of sensor networks one would include SSN interested features extending it with units, location, feature and time ontologies [4]. Additional classes and properties can be defined and added to model domain specific knowledge.

There are also initiatives dealing with sensor streaming data on query level. In [26], Shahriar et al. propose a smart query system considering both streaming data and historical data from marine sensor networks. ES3N [18] and C-SPARQL [23] are also dealing with sensor stream data. C-SPARQL is an extension of SPARQL for supporting stream data querying. Query processing is an important issue on the Semantic Sensor Web [26], but it is out of the scope of this paper. Instead, we focus on rule layer reasoning.

3 Rule-Based Reasoning

As claimed in the previous section almost every sensor network knowledge base is modeled through OWL ontologies. The Semantic Sensor Web foundation has enabled semantic enrichment of simple sensor data through these ontologies. However, inferring new and implicit knowledge from known facts represented in ontological terms is enabled through a powerful mechanism known as rule-based reasoning. In general, the limited expressivity of the Sematic Web Rule Layer (SWRL) [26], which currently has the status of W3C submission, has forced the community to consider hybrid systems while keeping the knowledge base modeled in the form of ontologies. Specifically, for the domain of sensor data an obstacle appears from the continuous flow of data. These data need to be consumed quickly by the reasoning engine which

⁵ http://www.loa-cnr.it/ontologies/DUL.owl

in turn will efficiently infer new knowledge by combining these data with background knowledge. Because of this nature when trying to infer logical consequences from sensor data, different rule systems are considered by the community. In general, rule systems fall into three categories: first-order, logic programming (LP), and action rules [29]. In the rest of this section we will briefly describe the main rule systems developed for modeling water quality management systems.

3.1 Association Rule Mining

Association rule mining is about finding frequent paths and correlations between items in the database. In [8], Ding et al. have proposed a framework for association rule mining and scoping in spatial datasets [8]. For example, they have used an association rule to infer dangerous arsenic levels with 100% confidence.

As envisioned by Bhatnagar and Kochhar [7], association rule mining performing on stream data are increasingly in need. They are employed in the estimation of missing data streams of data generated by sensors and frequency estimation of internet packet streams [7].

Association rule mining is more concerned with predicting what may happen in the future, while our aim is to deal with the current state of water quality.

3.2 Production Rules

Production rules are IF-THEN rules which fire actions (the THEN part) based on the precondition (the IF part) matching the current "state of the world". In [27], the authors model a hybrid Environmental Decision Support System (EDSS) for Waste-Water Treatment Plants (WWTP). As an example of production rules they infer invalid NO3 measurement values. They argue that the WWTP domain should be modeled through ontologies, for modeling sensor data, paired with decision-making rules, for processing incoming sensor data and recommending actions to be taken.

Another system implemented in terms of production rules has been designed by Chau [13] in the domain of water quality modeling. Namely, the system simulates human expertise during the problem solving of coastal hydraulic and transport processes. Both forward-chaining and backward-chaining are used collectively during the inference process [13].

The W3C in 2005 has created the Rule Interchange Format (RIF)⁶, a standard for rule exchange, with one of its dialects RIF-PRD⁷ implementing production rules. RIF-PRD enables interoperability of different rule systems with Semantic Web standards [29].

As proved in [27] production rules can be efficiently employed on a hybrid model for dealing with sensor networks. Production rules can modify and retract knowledge base facts, but they lack a protocol for wide publication.

⁶ http://www.w3.org/TR/rif-overview/

⁷ http://www.w3.org/TR/2013/REC-rif-prd-20130205/

3.3 Event Processing

Another hybrid approach which deals with sensor data reasoning is using OWL ontologies and CEP, which is an area similar to Stream Reasoning. CEP provides on-the-fly real-time processing of streams of atomic events (e.g. sensor readings) [14], [15]. Taylor and Leidinger [16] translate the whole OWL ontology, which models the event definition and optimization and extends an early version of the SSN ontology, into CEP statements for processing in an event processing engine. Unlike this approach, Anicic et al. [15] have taken the advantage of both "worlds" synthesizing the ability of CEP systems to process real-time complex events within multiple streams of atomic occurrences and the Semantic Web i.e. ontological ability to effectively handle background knowledge and perform reasoning. The later approach has resulted in a new rule-based language ETALIS [14] and EP-SPARQL [15], a query language extending the SPARQL language with event processing and stream reasoning capabilities. Both are implemented in Prolog, which has its foundations in LP.

All in all, CEP systems are very useful while dealing with real-time detection of sequential events e.g. if X water quality parameter drops from its allowed threshold, then another Y water quality parameter should be measured and if it is lower (or greater) than some constant *a*, than the river will be classified as 'polluted' and a water expert should be alerted. However, CEP cannot be used as a standalone system for reasoning in sensor networks. In fact, they lack the reasoning module, which is complemented by Semantic Web as it is the case with ETALIS [14].

3.4 Semantic Web

Dealing with sensor data, a pure Semantic Web approach has been implemented by Keßler et al. [5]. They have leveraged the SWRL's ability to express free variables and the use of its built-ins for modeling mathematical functions which has fulfilled the OWL's lack of mathematical processing capabilities. The approach is tested for geographical information retrieval (GIR) task for recommending personalized surf spots based on user location and preferences. A similar approach is taken by Wei and Barnaghi [6] who demonstrate how rule-based reasoning can be performed over sensor observation and measurement data within the terms of Semantic Sensor Web. They emphasize the ability of rules not just to infer accurate but also approximate knowledge.

Approaches [5] and [6] have proved that expressing SWRL rules in the domain of sensor data is feasible. However, they lack implementation and do not deal with reasoning obstacles i.e. monotonicity and open world assumption and their implications. These will be addressed in more detail in the next section.

Henson et al. [21] have used the Jena Semantic Web Framework [12] as an engine for reasoning with rules implemented for Semantic Sensor Web on weather domain. Using Jena rules they infer new knowledge about sensor observation data and link the newly generated relations with original observation time and location data. This approach suffers from the monotonicity issue, since Jena supports monotonic inference by default.

4 Discussion

Semantic Sensor Web enables semantic sharing and reasoning over sensor data spread all over the world. Inferring new knowledge from these data represents a huge improvement. However, this achievement cannot be easily obtained, because of the streaming nature of sensor data. During our previous works in [1] and [2], attempting to implement a database normalization tool through Semantic Web technologies we have encountered different issues on the ontological and rule level. OWL and SWRL's open world assumption and monotonic reasoning implied difficulties while dealing with negation as failure, classical negation, knowledge base modifications and disjunction. We think we will encounter the same reasoning challenges as sensor networks community is committed to Semantic Web, and Description Logic (DL), on which are based Semantic Web formal foundations, is the one to blame for this. Because of this inadequacy of DL, the preferred implementation has often been to use LP instead. As an example, ETALIS [14] is a rule-based system which reasons over sensor data and it is implemented in Prolog, which is a LP language.

4.1 Streaming Nature of Sensor Data

Wei and Barnaghi [6] and Keßler et al. [5] represent how SWRL reasoning can be effectively leveraged in Semantic Sensor Web. Henson et al. [21] use Jena rules [12] to infer new knowledge from sensor observation data. All these approaches do not consider the streaming nature of sensor data and thus the continual modification of sensor outputs. As envisioned by Barbieri et al. [28], because Semantic Web is still focusing on static data the continuous processing of data streams together with rich background knowledge, requires specialized reasoners. As a result, they have developed C-SPARQL to enable querying over stream data in combination with static background knowledge. We envision a need for a rule system which will deal with rapid flow of sensor data in combination with static knowledge. Perhaps a similar approach to Barbieri et al. [23] should be considered by SWRL developers to enable continuous execution of rules over sensor data.

4.2 Open World Assumption and Negation As Failure

The open world assumption deals with the completeness of the knowledge about a particular domain. In contrast to closed world assumption where the absence of a fact will return false, in open world semantics what is not known is undefined. For example, one cannot be sure that a particular measurement site is not polluted just because the water quality parameters are within their allowed thresholds. If a property stating that a measurement site is not polluted is absent then it does not mean that this

property does not exist. This uncertainty makes it difficult to consider pure Semantic Web approaches. As a consequence of open world assumption we cannot assume the negation as failure. Negation as failure means that we can infer not P if we fail to prove every possible proof of P. For example, water temperature should not reach the rates greater than 1.5° C plus ambient temperature for general conditions. Expressing this rule is not possible in DL, because of not supporting the negation as failure.

LP and production rules approaches have the advantages of supporting closed world assumption with negation as failure. In this manner their knowledge is defined. This is why the sensor networks community has implemented a hybrid Semantic Web approach, layering closed world reasoning on top of open world assumption i.e. OWL. In general, CEP is attractive, but we do not have any basis for confidence in how it works within a system making mixed assumptions about open and closed worlds. A possible approach is using query languages adopting closed world assumption, which are out of the scope of this discussion. Instead, we are focused on rule-based reasoning.

4.3 Monotonicity

OWL and SWRL support monotonic inference only i.e. earlier conclusions cannot be invalidated. Dealing with sensor data, we agree with Keßler et al. [5], who state that the same reasoning steps may lead to different results, thus the problem of monotonicity need to be addressed.

Let's consider the following example:

A sensor node, which is a collection of sensing devices (e.g. sensor probes) placed in an observation point of the river, will report values about water quality parameters which are within the allowed threshold. So, the sensor node will be classified as "clean". Suddenly, a parameter drops from its allowed threshold thus the sensor node should be classified as "polluted".

Because of the monotonic inference, the sensor node firstly asserted as "clean" cannot be later modified as "polluted". Additionally, as indicated in [22] the windows opened over streams can determine changes in the static information sources. This inadequacy has forced the sensor networks community to enable nonmonotonic reasoning through hybrid approaches described in the previous section. They have layered LP e.g. Prolog system, or production rule systems reasoning with updated working memory, to enable nonmonotonic inference.

5 Conclusion

In this paper we described current trends on representing sensor data with semantic technologies and their pros and cons. Because of the streaming nature of sensor data the community has mainly taken two approaches: building a hybrid system or a pure Semantic Web system. Both approaches prove that the Semantic Web rule layer lacks the capabilities needed to efficiently reason over sensor data. This is due to the issues of open world semantics and monotonicity. With some examples from the domain of

water quality management we have demonstrated the inability of a pure Semantic Web approach. Because of this shortage the sensor networks community has integrated ontologies with closed world and nonmonotonic reasoning.

Acknowledgements. The work described in this paper was supported by "InWaterSense: Intelligent Wireless Sensor Networks for Monitoring Surface Water Quality", an EU funded project managed by European Union Office in Kosovo, implemented by University of Prishtina.

References

- Ahmedi, L., Jajaga, E.: Normalization of relations and ontologies, In: The 10th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Databases, pp. 419-425, Cambridge, UK (2011)
- Ahmedi, L., Jajaga, E.: A database normalization tool using Semantic Web technologies, In: International Journal of Systems Applications, Engineering and Development, vol. 5 (2011)
- 3. Sheth, A., Henson, C., Sahoo, S.: Semantic Sensor Web, In: IEEE Internet Computing, pp. 78-83 (2008)
- 4. Compton, M., Barnaghi, P., Bermudez, L., Garcia-Castro, R., Corcho, O., Cox, S., Graybeal, J., M. Hauswirth, C. Henson, A. Herzog, V. Huang, K. Janowicz, W. D. Kelsey, D. Le Phuoc, Lefort, L., Leggieri, M., Neuhaus, H., Nikolov, A., Page, K., Passant, A., Sheth, A., Taylor, K.: The SSN Ontology of the W3C Semantic Sensor Network, Incubator Group, Journal of Web Semantics (2012)
- Keßler, C., Raubal, M., Wosniok, C.: Semantic Rules for Context-Aware Geographical Information Retrieval, In: Barnaghi, P. (eds.) European Conference on Smart Sensing and Context, (EuroSSC 2009), LNCS, vol. 5741, pp. 77–92, Springer (2009)
- Wei, W., Barnaghi, P.: Semantic Annotation and Reasoning for Sensor Data, In: Smart Sensing and Context, pp.66-76 (2009)
- 7. Bhatnagar, V., Kochhar, S.: Association Rule Mining, Encyclopedia of Artificial Intelligence, pp. 172-178, IGI Global (2009)
- Ding, W., Eick, C. F., Yuan, X., Wang, J., Nicot, J.: A Framework for Regional Association Rule Mining and Scoping in Spatial Datasets, Geoinformatica, vol. 15, issue 1, pp. 1-28 (2011)
- 9. Ceccaroni, L., Cortes, U., Sanchez-Marre, M.: OntoWEDSS: an ontology-underpinned decision-support system for wastewater management (2001)
- Ceccaroni, L., Cortes, U., Sanchez-Marre, M.: WaWO-An ontology embedded into an environmental decision-support system for wastewater treatment plant management, In: Proceedings of ECAI2000 - Wo9: Applications of ontologies and problem-solving methods, pp. 2.1-2.9, Berlin, Germany (2000)
- 11. Hitzler, P., Krotzsch, M., Parsia, B., Patel-Schneider, P., Rudolph, S., OWL 2 Web Ontology Language Primer, http://www.w3.org/TR/owl2-primer/ (2009)
- 12. Jena Semantic Web Framework, http://jena.sourceforge.net/, last accessed July, 2nd 2013
- 13. Chau, K. W.: An ontology-based knowledge management system for flow and water quality modeling, In: Advances in Engineering Software, vol. 38, no. 3, pp. 172-181 (2007)
- 14. Anicic, D., Fodor, P., Rudolph, S., Stuhmer, R., Stojanovic, N., Studer, R.: A Rule-Based Language for Complex Event Processing Reasoning, In: Proceedings of the Fourth

International Conference on Web reasoning and rule systems, pp. 42-57, Springer-Verlag Berlin, Heidelberg (2010)

- Anicic, D., Fodor, P., Rudolph, S., Stojanovic, N.: EP-SPARQL: A Unified Language for Event Processing and Stream Reasoning, In: WWW 2011, pp. 635-644 (2011)
- Taylor, K., Leidinger, L.: Ontology-Driven Complex Event Processing, In: ESWC, LNCS, pp. 285-299, Springer, Greece (2011)
- 17. Amato, G., Caruso, A., Chessa, S., Massi, V., Urpi, A.: State of the Art and Future Directions in Wireless Sensor Network's Data Management, In: Istituto di Scienza e Tecnologie dell'Informazione del CNR, (2004)
- Lewis, M., Cameron, D., Xie, S., Arpinar, I. B.: ES3N: A Semantic Approach to Data Management in Sensor Networks, In: A workshop of the 5th International Semantic Web Conference ISWC (2006)
- Gruber, T. R.: A translation approach to portable ontologies, In: Knowledge Acquisition, vol. 5(2), pp. 199-220, (1993)
- Barbieri, D., Braga, D., Ceri, S., Della Valle, E., Huang, Y., Tresp, V., Rettinger, A., Wermser, H.: Deductive and Inductive Stream Reasoning for Semantic Social Media Analytics. In: Proceedings of the Seventh Extended Semantic Web Conference (ESWC'10), L. Aroyo, G. Antoniou, E. Hyvonen, A. Teije, H. Stuckenschmidt, L. Cabral, and T. Tudorache, Eds. LNCS, vol. 6088, pp. 1–15. Springer-Verlag (2010)
- Henson, C. A., Pschorr, J. K., Sheth, A. P., Thirunarayan, K.: SemSOS: Semantic Sensor Observation Service, In: Proceedings of the 2009 International Symposium on Collaborative Technologies and Systems (CTS 2009), Baltimore, MD (2009)
- Della Valle, E., Ceri, S., van Harmelen, F., Fensel, D.: It's a Streaming World! Reasoning upon Rapidly Changing Information, In: IEEE Intelligent Systems, vol. 24(6), pp. 83–89 (2009)
- Barbieri, D. F., Braga, D., Ceri, S., Della Valle, E., Grossniklaus, M.: C-SPARQL: SPARQL for Continuous Querying. In: Proceedings of the Eighteenth International Conference on World Wide Web (WWW'09), Quemada, J., Leon, G., Maarek, Y., Nejdl, W. (eds.), pp. 1061–1062, ACM Press (2009)
- Bolles, A., Grawunder, M., Jacobi, J.: Streaming SPARQL extending SPARQL to process data streams. In: Bechhofer, S., Hauswirth, M., Hoffmann, J., Koubarakis, M. (eds.) ESWC 2008. LNCS, vol. 5021, pp. 448–462. Springer, Heidelberg (2008)
- Rodriguez, A., McGrath, R. E., Liu, Y., Myers, J.: Semantic Management of Streaming Data. In: Proceedings of International Workshop on Semantic Sensor Networks, SSN, pp.80-85 (2009)
- Shahriar, M. S., De Souza, P., Timms, G.: Smart Query Answering for Marine Sensor Data, Sensors 2011, vol. 11(3), pp. 2885-2897 (2011)
- 26. Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosof, B., Dean, M.: SWRL: A Semantic Web Rule Language Combining OWL and RuleM, W3C Member Submission, http://www.w3.org/Submission/SWRL/, last accessed April, 22nd 2013 (2004)
- 27. Sottara, D., Bragaglia, S., Mello, P., Pulcini, D., Luccarini, L., Giunchi, D.: Ontologies, Rules, Workflow and Predictive Models: Knowledge Assets for an EDSS, In: Seppelt, R., Voinov, A. A., Lange, S., Bankamp, D. (eds.), International Environmental Modelling and Software Society (iEMSs), 2012 International Congress on Environmental Modelling and Software Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Germany (2012)
- 28. Barbieri, D., Braga, D., Ceri, S., Della Valle, E., Grossniklaus, M.: Stream Reasoning: Where We Got So Far. In: Proceedings of the 4th International Workshop on New Forms of Reasoning for the Semantic Web: Scalable and Dynamic (NeFoRS) (2010)

- 29. Kifer, M., Boley, H.: RIF Overview (Second Edition), W3C Working Group Note 5 February 2013, http://www.w3.org/TR/rif-overview/, last accessed July, 2nd 2013
- 30. Schaaf, M., Grivas, S. G., Ackermann, D., Diekmann, A., Koschel, A., Astrova, I.: Semantic Complex Event Processing, In: Proceedings of the 5th WSEAS congress on Applied Computing conference, and Proceedings of the 1st international conference on Biologically Inspired Computation, pp.38-43 (2012)

Enrichment of Digital Libraries with Web 2.0 Resources for Enhanced User Search Experience

Fidan Limani¹, Vladimir Radevski¹

¹South East European University, Faculty of Contemporary Sciences and Technologies, Bul. Ilindenska, n. 335, 1200 Tetovo, Republic of Macedonia {f.limani, v.radevski}@seeu.edu.mk

Abstract. The publications stored in Digital Libraries (DL), while authoritative and of high quality, due to DL's access model, have limited chances of being linked up with other, up-to-date resources on the Web.On the other hand, the content amassing in Web 2.0-driven resources in the form of scientific blogs and wikis is continually on the rise. Although with no formal publication "check points" or review committees, their key feature – the up-to-datedness – both in the form of ease of publication, and the opportunity for feedback from the community that follows these resources – is something that Digital Libraries obviously lack. With this proposal, by using semantic web and Linked Data technologies, we want to develop a service that will enrich the user search experience by combining the content from DLs with that of Web 2.0-driven scientific blogs and wikis.

Keywords: Semantic Web; Scientific blogs and wikis; Enriching and enhancing search.

1 Introduction

The Semantic Web initiative aims at extending the World Wide Web by "enriching" the data on the Web and rendering it machine-processable, which, in turn, lessens the burden of human involvement in processing the data, be it in search–, integration–, or mining – related operations, to name a few. Furthermore, through the Linked (Open) Data project, any data provider is encouraged and able to publish their enriched or structured data sets online so that the potential data consumers on the Web can link up and make use of all the "attractive" available data sets of this nature.

Semantic web technologies are already being applied in improving applications in many domains, including that of Digital Libraries. The Library Linked Data Incubator Group [7] is leading the efforts towards building a common infrastructure supported by Semantic Web standards, which will enable collaboration and content sharing – primarily between the library communities – spanning other actors from broader domains, such as archives, museums, publishing, the Web, etc.

As shown in [1], semantic web technologies, by enabling data integration from multiple, often heterogeneous (contributions from library user communities in the form of resource annotations), resources, can be used to increase the efficiency of information discovery in Digital Libraries.

On par with the user-generated metadata generation, which can be quite helpful in indexing and maintaining metadata for DL-s, there are automatic services for automatic metadata creation. One research in particular, [2], proposes a technique for evaluating the semantic techniques applied by these services. While the information availability in the form of researcher profiles and publications is abundant, the focus of one research publication [6] finds that contextualizing these profiles and their social network activities to suggest potential research collaborations would bring great value to the participants of the research community.

As more and more resources of interest are accruing on the Web, users will need to efficiently sift through this information and deal with the information overload that they face frequently during the Web-related activities [3], while having to put in the minimum effort possible in the process. In particular, resources generated by the Web 2.0-driven scientific blogs and wikis present an important resource candidate for linking up to the DL publication pools. As more and more researchers use this channel as their publication tool of choice, the need to interlink these data repositories with the DL's list of publications becomes quite an interesting use case to treat, and also the focus of this proposal.

2 Problem Statement

One of the contributions of the Semantic Web initiative is the enrichment of the search experience through data integration. The applications in this domain would be able to merge, or integrate, several resources based on the semantics encoded along the data. As a result, if the concept we are searching for spans over several disciplines or, in our case, through several Web publication channels, the proposed approach should be able to query all the related terms regardless of their indexation.

There already are studies showcasing the potential of data integration through the Semantic Web technologies, with focus on scholarly research and collaboration. One study [3] provides good results from semantic-based researcher recommendation services able to match researchers that can potentially collaborate in future projects for mutual gains. In a similar approach, another study, [6], treats the identification and interlinking of different research groups for building collaborative relationships, again based on services supported by Semantic Web technologies.

The scenarios of interest for this research proposal are from the domain of Digital Libraries, and focus on the inclusion of Web 2.0-generated data – Scientific Blogs and Wikis, as well as the Linked Open Data as a means for publishing structured data, as well as enabling data from different sources to be interlinked and queried.

Scientific blogs and wikis are interesting development in scientific publishing, commenting, peer reviewing, and sharing, that is gaining continuous traction. Writing scientific blogs departs from the traditional research channels. However, scientists find blogs and wikis as appropriate channels for many reasons: disseminating their findings is faster and more convenient; the feedback they get from the community is faster; they share the information on their experiments - both the successful and unsuccessful ones (which rarely get reported on in a research paper or journal, for example), and many other benefits.

The aim of our research is to integrate "traditional" scientific research resources – properly indexed, categorized, and searchable; with the scientific Web 2.0 data, or socially maintained scholarly libraries services, regardless of the quality of data provided in these Web 2.0-backed resources. Scientific blogs and wikis allow users to tag resources, information that can later be used to link these sites with "authoritative" (library-supported, professionally-catalogued, indexed resources) scientific literature, with the goal of enriching these resources. The richness that the users would provide while visiting these blogs or wikis is welcome, of course: they would be able to easily tag the contents in those resources (through tags, comments, and sharing count, for example), thus adding more qualifiers to the data presented there, and, as a result, increase the results and the relevance of the data integration.

A model scenario to be considered is the following: A researcher studying the effect of how technology innovation affects the national Gross Domestic Product, for example, would firstly search for resources hosted on "authoritative" institutions, such as (digital) libraries, research institutions, or specialized resources that focus on the concepts of interest. However, there may be important information in the form of a scientific blog or wiki article, where a governmental representative, such as a Minister of Technology or Economy, or an expert on the matter, would express their views on various issues and latest developments regarding - in this case - the relationship between the technology innovation and its effects on the national GDP. This, without a doubt, would be very relevant information to the researcher. Being a scientific 2.0 resource, its users will provide additional information about the article. They could use various categories to tag the information on the said article; provide comments on the article; share the article with their friends on social networks; thus amassing important data related to that scientific posting. In this way it would be possible to recommend or link the user to the most "appropriate" resources relevant to their research topic (the most commented; the most shared; the most read; etc) by mining this very information on the scientific 2.0 resources, resulting in the best search results for the user.

The scenario could be set in the reverse flow, too: the user, after visiting her/his favorite scientific blog or wiki, would "pick" an article of interest – technology innovation affecting the national GDP, in our case. After the initial search results, s/he can decide to further narrow her/his scope by considering additional factors, which are specific for the Web 2.0-driven scientific blogs and wikis, such as the most commented; the most shared; the most read; etc. If s/he wants to have a firm grasp on the topic, and consult a more "authoritative" publication, now that s/he has identified the author that interests her/him more, s/he can undertake a search in a Digital Library repository.

In this scenario, regardless of where the user starts off with her/his search, as demonstrated by this example, the result of interlinking data resources (in this case research publications) enriches and enhances the search results as the user is not only isolated in the realm of Digital Libraries publications, but can link up those resources with the publications that are part of scientific blogs and wikis publication pool.

3 Proposed Solution

The focus of our research proposal revolves around data sources in Digital Libraries and Scientific Blogs and Wikis. More specifically, the target data repositories for our research proposal are EconStor¹ and EconBiz² as representative data stores of the former, and publications from a list of scientific blogs and wikis from the domain of economy³, as representatives of the latter.

We are proposing a service that, initially, will match publication resources in these two different data repositories "types"; potentially provide new information on publications that could be of interest to the user or are somehow (semantically) related to the initial search without the user specifically requesting or knowing about it beforehand; as well as suggest new publications that might be of interest to the user – a recommendation service for scholarly publications.

This service will (a) help enrich the content of both DLs and scientific blogs and wikis by the mere fact that the user would seamlessly pool for publications in both repository types; (b) enable crossing the DL "boundary" by interlinking the different publications in both repositories types; (c) enhance the user search experience as the user will, ultimately, receive not only the added value from considering the two publication repositories, but also receive new insights and recommendations for new articles to consider; (d) allow users to engage more to the content as, especially on the side of scientific blogs and wikis – users will be able to comment, tag, share, etc, the publications of their interest; (e) bring more freedom to the users as they will have more choice in searching for publications of their interest.

If we carry out a search in the EconStor data repository on the topic of IT innovation affecting national GDP, the service we are proposing should offer us more results than before – benefit (a); more qualitative and up-to-date results in terms of publications, comments, and other insightful information – benefit (b); due to the inferencing component of the service, the user will "stumble" upon an interesting and related publication which s/he did not though of before – benefit (c); after selecting a publication of interest, tag it according to their understanding, share it with her/his friends, comment it to express her/his view, etc – benefit (d); in some cases, just following their author of interest – which they either "discovered" through the initial search or the service recommended it – in the scientific blogs and wikis publications will be enough to keep up to date with the developments in the domain of interest to the user – benefit (e).

¹ Open Access server of the German National Library of Economics – Leibniz Information Centre for Economics.

² EconBiz is a joint project of the German National Library of Economics (ZBW) and the University and City Library of Cologne (USB Köln).

³ The list of scientific blogs and wikis to interlink with the Digital Library platforms EconStor and EconBiz is determined by the German National Library of Economics (ZBW).

4 Plan

We envision developing the service in four phases. The *first* phase should automate the navigational-type search, which follows the path outlined by the various relationships to find data that users can perform even now, but will have to manually navigate through the terms or concepts of interest to them. The *second* phase focuses on making suggestions to the user by carrying out reasoning-related operations, or trying new set of rules on the combined data from both repositories (EconStor/EconBiz and scientific blogs and wikis). The *third* phase is oriented towards turning the suggestions from the previous two phases into a personalized library for the user. Ideally, the library will consist of all the publications that the user searched for (and found them as a result of the search), as well as the publications that were suggested to them from the systems (as a result of the Phase 1 and 2), and also were accepted by the user. One final phase of the research, the *fourth*, entails the user evaluation. During this phase a user base will evaluate their search experience related to the system. Based on the results from this phase, the readiness and success of the system could be determined.

5 Related Work

In this article we describe the benefits of developing a service that will interlink resources in DL repositories and Web 2.0-driven Scientific Blogs and Wikis – a service that will help enrich enhance the user experience from using DL and Scientific Blogs and Wikis resources.

As demonstrated in [1], the user experience (efficiency increase of information discovery in Digital Libraries) could be improved by enabling data integration from multiple resources, such as contributions from library user communities in the form of resource annotations, for example.

Two other studies, [3, 6], focus on exploiting researcher profiles and publications and their social network activities to suggest new matching possibilities between researchers.

With the same goal in mind as the research in [3] and [6], the PIRATES framework [5] extends the access type to a DL by including access to publications semantics with the goal of enhancing the user experience while using DL-s – the E-Dvara digital library platform in particular. End users' search is extended to include not only the usual search queries, but user profiles are also used when processing the user requests.

Sun J. et al. propose a novel approach for the personalized article recommendation in online scientific communities [8]. This research contribution is especially in relation to the Phase 1 and 2 from the roadmap, as represented in figure 1.

6 Conclusion

The publications stored in Digital Libraries (DL), while authoritative and of high quality, due to DL's access model, have limited chances of being linked up with other, up-to-date, resources on the Web.

On the other hand, the content amassing in Web 2.0-driven resources in the form of scientific blogs and wikis is continually on the rise. Although with no formal publication "check points" or review committees, their key feature – the up-to-datedness – both in the form of ease of publication, and the opportunity for feedback from the community that follows these resources – is something that Digital Libraries obviously lack.

With this proposal, by using semantic web and Linked Data technologies, we want to develop a service that will enrich the user search experience by combining the content from DLs with that of Web 2.0-driven scientific blogs and wikis.

There is a need for intelligent services which can link the authoritative publication resources from Digital Libraries to the most recent content from scientific blogs and wikis to the user automatically by reducing their search span and effort. With this proposal, by using semantic and linked data technologies, we want to develop a service that will enrich the user search experience by combining the content from scientific libraries with that of Web 2.0 blogs and wikis.

Acknowledgements

This material is based upon works supported by the German National Library of Economics (ZBW), under the mentorship of Prof. Dr. Klaus Tochtermann, and assistantship of Dr. Atif Latif.

7 References

- 1. Kruk, S. R.: Semantic Digital Libraries: Improving Usability, Springer (2010)
- Tönnies, S. and Balke, W-T.: Using Semantic Technologies in Digital Libraries A Roadmap to Quality Evaluation, Research and Advanced Technology for Digital Libraries, Lecture Notes in Computer Science. Vol. 5714/2009, p. 168-179; (2009)
- 3. Y. Xu, J. Hao, R. Y. Lau, J. Ma, W. Xu, and D. Zhao: A personalized researcher recommendation approach in academic contexts: Combining social networks and semantic concepts analysis. PACIS (2010)
- Strohmaier M., Körner Ch., and Kernb R.: Understanding why users tag: A survey of tagging motivation literature and results from an empirical study; Web Semantics: Science, Services and Agents on the World Wide Web, vol. 17 p. 1–11. Elsevier (2012)
- Baruzzo, A., Casoto, P., Challapalli, P., Dattolo, A., Pudota, N., Tasso, C.: Toward Semantic Digital Libraries: Exploiting Web2.0 and Semantic Services in Cultural Heritage, Journal of Digital Information, vol. 10(6), p. 4. (2009)
- 6. Morales-del-Castillo J. M., Peis E., and Herrera-Viedma E.: An Approach to a Semantic Recommender System for Digital Libraries. Semantic Web Personalization and Context

Awareness: Management of Personal Identities and Social Networking (pp. 55-68). Hershey, PA (2011)

- Library Linked Data Incubator Group Charter, <u>http://www.w3.org/2005/Incubator/Ild/charter</u>
- 8. Sun J. et al.: A Novel Approach for Personalized Article Recommendation in Online Scientific Communities. 46th Hawaii International Conference on System Sciences (2013)

Towards Continuous Quality Assurance in Future Enterprise Cloud Service Brokers

Dimitrios Kourtesis^{1,2} and Konstantinos Bratanis^{1,2},

 ¹ South-East European Research Centre, The University of Sheffield,
 24 Proxenou Koromila, 54622, Thessaloniki, Greece, {dkourtesis, kobratanis}@seerc.org

² Department of Computer Science, The University of Sheffield, Regent Court 211 Portobello Street Sheffield S1 4DP, UK, {d.kourtesis, k.bratanis}@dcs.shef.ac.uk

Abstract. Cloud service brokerage is an emerging operational model in the context of cloud computing. A cloud service broker acts as an intermediary between a service provider and a service consumer using various brokerage capabilities, in order to facilitate the provision and the consumption of cloud services. Continuous quality assurance is a brokerage capability of high value to providers and consumers of cloud services, but it is also among the most challenging types of functionality for cloud service brokers to implement. In this paper, we focus on a subset of such brokerage capabilities for continuous quality assurance of cloud services. In particular, we focus on research challenges in the scope of cloud service governance and quality control, as well as cloud service failure prevention and recovery.

Keywords: Cloud computing, cloud service brokerage, cloud service governance, cloud service failure prevention and recovery, policy compliance, service-level agreement monitoring.

1 Introduction

With the increasing adoption of cloud computing the enterprise IT environment is progressively transformed into a matrix of interwoven infrastructure, platform and application services, delivered from diverse service providers. As the number of cloud service providers grows and the requirements of consumers become more complex, the need for actors to assume the role of an intermediary between providers and consumers is becoming stronger. Cloud service intermediation is increasingly recognized as an indispensable component of the cloud computing value chain.

Examples of existing cloud service intermediation offerings include services helping enterprises to find and compare cloud services (e.g. marketplaces/stores), to develop and customize services (e.g. aPaaS - application platform as a service offerings), to integrate services (e.g. iPaaS - integration platform as a service), to

monitor and manage services, and many more. Despite differences with respect to the capabilities such cloud service intermediaries offer, or how these capabilities are combined, they have one thing in common: making it easier, safer and more productive for cloud computing adopters to navigate, integrate, consume, extend and maintain cloud services. According to Gartner, this is precisely the value proposition of a 'Cloud Services Brokerage', a term coined in 2010 to refer to the emerging role of brokers in the context of cloud computing [1].

In the future, enterprises will demand brokerage capabilities much more sophisticated than what is on offer by cloud service intermediaries today. Continuous quality assurance of cloud services is one such type of brokerage capability; foreseen to be most valuable for service consumers, but at the same time most challenging for future brokers to implement.

In this paper we try to briefly introduce the concept of Cloud Service Brokerage (CSB) and to motivate the need for continuous quality assurance as an important intermediation capability of future enterprise cloud service brokers. We focus our attention on two specific forms of continuous quality assurance intermediation: (i) cloud service governance and quality control; and (ii) cloud service failure prevention and recovery. For each area we provide an overview of related work, and report on the research themes explored by the authors in a joint research effort under FP7 project Broker@Cloud.

2 Cloud service brokerage

As the number of externally-sourced services in an enterprise cloud environment increases, it becomes increasingly more difficult to keep track of when and how services evolve over time—be that through changes that providers make to the terms of provision or to the actual services, or through changes which are unintentional, such as variations in service performance and availability. Moreover, it becomes increasingly more difficult to appreciate all the different kinds of impact that a change can have with respect to a service's compliance to different policies and regulations, its conformance to normative technical specifications and contracts (Service Level Agreements), and generally, with respect to the fulfilment of all different kinds of functional and non-functional requirements surrounding a particular service's usage.

Cloud Service Brokerage (CSB) represents a new type of service and emerging business model in the space of cloud computing, aimed at helping enterprises to address precisely such challenges, and to mitigate the risks that ensue from the complexity inherent in large-scale enterprise cloud environments [2]. In an analogy to the way other kinds of intermediaries operate within different areas of traditional commerce, a Cloud Service Brokerage is an entity that works on behalf of a consumer of cloud services to intermediate and to add value to the services being consumed.

Much of the enabling technology that is needed to support different cloud service brokerage capabilities is certainly not new. Recent years have seen a proliferation of many relevant proprietary and open source tools that could provide fragments for the implementation of such services, such as tools for monitoring and managing applications, services and virtualized infrastructures, or tools for integrating heterogeneous data, processes and applications, to name a few. Companies that have already created cloud service offerings based on such enabling technologies (such as RentTheCloud, SpotCloud, CloudKick, Cloud Switch, Kaavo, Rightscale, Vordel, Zimory, JitterBit, SnapLogic, Boomi - acquired by Dell, or Cast Iron -acquired by IBM), can be considered as elementary, albeit primitive, forms of cloud service brokers.

3 Continuous Quality Assurance Intermediation

As future enterprise cloud service brokers will start to turn their attention to monitoring the obligations and opportunities pertaining to the delivery of cloud services to consumers, continuous quality assurance intermediation capabilities will become more and more important. Intermediation will cover software-based services ranging from simple programmatically-accessible web APIs, to complex software applications delivered as cloud services, i.e. on-demand Software-as-a-Service offerings.

Continuous quality assurance intermediation capabilities represent open research topics which, to the best of our knowledge, have only recently come to the attention of research communities working on cloud computing, service-oriented computing, business computing, autonomic computing, or of relevant subcommunities (e.g. service governance and policy compliance, software testing and verification, self-adaptive service-based systems, recommender systems, quality of service optimization). At the time of this writing, the theoretical and pragmatic challenges pertaining to the introduction of continuous quality assurance functions in brokers of software -based cloud services remain largely unexplored.

In the next two subsections we briefly introduce two particular forms of continuous quality assurance intermediation: (a) service governance and quality control; and (b) service failure prevention and recovery. A research roadmap that considers additional forms of intermediation capability in the broader scope of continuous quality assurance and optimisation can be found in [3].

3.1 Cloud service governance and quality control

The need for governance and quality control in the provision and consumption sides of services is naturally carrying over to cloud intermediaries, who are positioned in the middle of the service supply chain and act as consumers and providers of services at the same time.

Governance refers to policy-based management of the processes used to provision and consume cloud services, and policy-based management of quality properties of cloud services and service-related artefacts. More specifically, we can distinguish between two forms of governance [4], which are complementary:

 Process governance – defining and enforcing policies to ensure that cloud services are developed, tested, deployed, provisioned or retired in accordance to a specified lifecycle model with explicit conditions for phase transitions.

 Artefact governance – defining and enforcing policies to ensure that artefacts associated to cloud services (such as technical interface descriptions or service-level agreements) are conformant to explicit technical or business constraints.

In process governance, policy enforcement amounts to ensuring that all services brokered by a cloud service intermediary proceed through a standard series of lifecycle phases, in a manner consistent with pre-specified transition criteria. In artefact governance, policy enforcement amounts to ensuring that all artefacts associated to those services are valid with respect to any prescribed constraints, irrespective of which lifecycle phase the corresponding services may be found in.

3.2 Cloud service failure prevention and recovery

The quality of enterprise cloud services cannot be assumed to be static after onboarding them onto an enterprise cloud service brokerage platform. This is owing to the fact that the cloud services are operated by third parties, who are continuously evolving the provided cloud services to enhance their feature set, an act that can introduce unexpected faults into the cloud services and ultimately lead to quality degradation and partial or total failure of the cloud service. Also, a cloud service brokerage environment each cloud service is part of the ecosystem, and new services may appear while others might no longer be performing as expected with respect to some quality criteria. Thus, the operator of the cloud service brokerage platform needs to identify alternatives during on boarding of a cloud services, but also continuously at run-time, and to detect when a change is needed, such that prevention or recovery from failures of cloud services takes place through triggering the necessary adaptation actions.

Failure prevention and recovery refers to the run-time monitoring and adaptation processes used for collecting information about the operation of a cloud service (e.g., QoS information related to a service-level agreement, resource consumption, etc.) and enacting some action (e.g., substituting a cloud service, offering the service at a lower price, etc.) to recover or prevent a QoS failure. More specifically, it includes two types of monitoring:

- Reactive monitoring, collecting and analyzing information in near real-time to detect failures that have occurred. This involves the calculation of some QoS indicators based on the monitored information and the evaluation of that indicators against some thresholds, which are defined using rules.
- Proactive monitoring, collecting and analyzing information in near real-time to predict potential QoS failures that will occur in the near future. This involves the curation of historical QoS information which is used to extract patterns of past QoS behaviour, and therefore being able to identify when a past pattern emerges during the current operation of a cloud service.

In the event that a failure is detected or predicted, some sort of reasoning is necessary for deciding the action required for recovering or preventing a failure.

4 Related Work

4.1 Related work on cloud service governance and quality control

A look at different cloud service intermediaries, such as providers of cloud application platforms (aPaaS), reveals different approaches and implementations of process and artefact governance. On the Intuit Partner Platform¹ development and deployment of cloud services proceeds through four prescribed phases, each of which is called 'a line of development'. The phases are development, quality assurance, staging, and publishing. Similarly, in the Heroku platform², add-ons (i.e. third-party services) advance through the phases of development, alpha, private beta, beta, and general availability. In Force.com³ the majority of quality checks on cloud service artefacts are associated with a particular phase towards the end of the development and deployment process, referred to as 'security review'—though the scope of the review carried out is actually much broader than security. Rollbase⁴ has defined a similar 'application approval' phase before the stage of deployment, during which all artefacts associated with a cloud service are reviewed against the platform provider's policies.

The industrial state of the art in policy management (policy creation, management and enforcement) follows closely on the evolution of tool suites supporting different aspects of governance (e.g. cataloguing and storage, service lifecycle management, dependency tracking, and of course, policy management). Those tool suites will typically come in some form of registry and repository system – a class of systems having its roots in solutions initially developed for SOA governance [5], [6], [7]. Vendors of today's governance registry and repository systems support different means by which policies can be encoded and enforced [8], [9], [10], [11]. As shown by a recent survey of methods for policy management in contemporary open source registry and repository systems [12], a major weakness in the state of the art is the lack of proper separation of concerns with regard to defining rules for governance and acting upon them. Policy definition and policy enforcement are entangled in the implementation of a single software component—the policy evaluation engine. For the most part, the rules that a policy comprises are encoded in an imperative manner, in the same programming language that the registry and repository system has been implemented, and as part of the same code that checks the data for violations. This can be shown to create many negative side effects.

Recent advances in the state of the art in policy-based systems management have tried to address similar shortcomings by insisting on the design of software applications in a way that certain rules are kept separate from the main program logic. Those rules are captured though policy-specification languages and consulted at runtime via an engine or reference monitor when user activity dictates to do so [13].

¹ http://ipp.intuit.com/

² http://www.heroku.com/

³ http://www.salesforce.com/platform/

⁴ http://www.rollbase.com/

Several works motivated by similar objectives have focused on the enhancement of existing policy languages and tools with ontology-based methods of representation and processing. The most prominent early works along this line were KAoS [14], Ponder [15], and Rei [16]. Other, more recent works in a similar direction are those by Kolovski et al [17] and Kolovski and Parsia [18].

Several recent research efforts and projects for industrial prototype development have been turning their attention to the benefits that the application of ontology-based modelling and reasoning can have with respect to different aspects of software engineering [19]. As Bergman points out [20], many of the benefits which are generally obtained by ontology-centric approaches to the development of information systems are attributed to the fact that the locus of effort is shifted from software development and maintenance to the creation and modification of knowledge structures. Uschold cites six important benefits which result from the increased level of abstraction and the use of formal structures and methods in ontology-driven information systems: reduced conceptual gap; increased automation; reduced development times; increased reliability; increased agility/flexibility; decreased maintenance costs [21].

The above benefits of ontology-centric approaches to information systems engineering, in combination with the Semantic Web standards and tools currently available [22], appear to provide a promising foundation for addressing the shortcomings of policy definition and enforcement in contemporary registry and repository systems [12]. Results from efforts with similar objectives are already being reported in the wider context of policy-driven systems management, such as the work by IBM on the transformation of sources of management data into Linked Data providers to allow for uniform logical queries over heterogeneous systems in a network [23].

4.2 Related work on cloud service failure prevention and recovery

Over the past decade there has been an increasing interest in incorporating selfmanaging capabilities in software systems, motivated by high complexity involved in the everyday administration, as well as, the detection, diagnosis, and resolution of failures in software systems. The research fields of Autonomic Computing [24] and Self-Adaptive Systems [25] continue to demonstrate fundamental advances towards understanding the challenges associated with the aforementioned research directions, while proposing fragmented approaches for partially addressing the reactive or proactive detection of failures, the diagnosis process to identify the cause of failure and, the automation of adaptation actions, as a solution to recovering from failures.

Recent research works, such as Bertolino et al. [26] and Romano et al. [27], are addressing the topic of cloud service monitoring at the infrastructure level, and they do not consider cloud services related to the platform and application levels. The literature addressing such issues in the context of cloud computing is almost non-existent and therefore it is necessary to adopt approaches from neighbouring fields, such as service-oriented computing, where such approaches have been elaborated more.

Research in self-adaptive service-based systems outlined in Di Nitto et al. [28] can serve as the foundation for exploring novel mechanisms for cloud service failure prevention and recovery in the context of cloud service brokerage, since cloud services, ranging from programmatically-accessible web APIs to complex software applications delivered as a service, possess principles similar to the ones defined by the service-oriented architectural style. Therefore, the research literature focusing on self-adaptive service-based systems is considered highly relevant.

Monitoring is well studied in self-adaptive service-based systems. Most monitoring approaches follow either the push mode, where events or date are sent to the monitoring component, or a pooling mode, where the monitoring component queries the subject of monitoring. The monitored information is often processed using an analysis approach, in order to detect an adaptation need. Such analysis approaches range from verification of service behaviour [29], and evaluating rules for detecting SLA violations [30], to dependency analysis for identifying causes offending some KPIs [31], and complex event processing for identifying situations or complex event based on the correlation of basic events [32].

The main challenge addressed by research in self-adaptive service-based systems is adaptation, i.e. the ability to reconfigure applications so as to support continuous, unimpeded augmentation of services in response to changing environmental circumstances. The notion of adaptation has been extensively studied in the computer science domain as it is considered as one of the most desired functionalities of today's highly dynamic, distributed and ubiquitous environments in the service-oriented setting [33]. Still yet, the realisation of such properties is limited to fragmented approaches and research prototypes exemplifying a particular aspect of the adaptation problem [34]. Adaptation, the process of modifying a system or application in order to satisfy new requirements and to fit new situations, can be performed either because monitoring has revealed a problem or because the application identifies possible optimizations or because its execution context has changed.

The main topic of works related to service adaptation, such as [35], [36], [37], focuses on QoS parameters (e.g., response time, failure rates, availability) and infrastructure characteristics (e.g., load, bandwidth) which are monitored and adjustments are made if there are deviations from some expected values. These approaches follow the reactive approach to adaptation, i.e., the modification of the application takes place after the critical event happened or a problem occurred.

There exist also few initial efforts for implementing proactive approaches such as the ones that systematically test services to uncover failures and deviations of the quality of services from the expected one. Existing approaches for testing service-based systems mostly focus on testing during design time, which is similar to testing of traditional software systems [38] but there are also others like PROSA [39] that exploit online testing solutions [40], at run-time in order to proactively trigger adaptations. In the same direction, regression testing functionality of that system but their main disadvantage is that they constrain it by considering the exchange of individual services as the only mechanism for performing adaptations of service-based applications (SBA). The aforementioned efforts only consider QoS parameters in adaptation; they cannot take into account 'external' issues related to the

environment of a service that shape specific situations, in which a system should be adapted.

Aspect-Oriented Programming (AOP) has been used in several research efforts, such as [41], [42], [43], [44], as a way to weave alternative actions in services that are part of business processes at run-time.

The high fragmentation of existing approaches has led to addressing only specific problems, particular perspectives or particular types of systems and often the monitoring solutions are isolated from the analysis of adaptation need, the reasoning of adaptation strategies, and the enactment of adaptation actions. Therefore, the isolated operation of monitoring and adaptation mechanisms are not adequate for the purposes of cloud service failure prevention and recovery in the context of an enterprise cloud service brokerage platform that requires the integrated operation of monitoring and adaptation mechanisms under a holistic solution, which can scale to deal with the increased load and to be extended to deal with the increased complexity.

Cross-layer monitoring & adaptation (M&A) has been acknowledged as promising solution for addressing the integrated operation of monitoring and adaptation mechanisms [45]. Cross-layer M&A aims at properly diagnosing the source of a problematic situation leading to identification of the effective adaptation requirements, while it also aims to generate a holistic adaptation strategy to be used for coordinating the adaptation actions across the layers of the managed element. Cross-layer M&A considers the relationships of system elements and the impact of different adaptation actions to different elements at separate layers of the system.

In order to employ cross-layer M&A for the purposes of cloud service failure prevention and recovery, several issues related to flexibility, extensibility, dynamicity and effectiveness need to be addressed [46], as well as the support for the combination and cross-correlation of observations, predictions and events from different sources and provided by different monitoring and adaptation mechanisms [45].

5 Conclusion & Future Work

Cloud service intermediation is becoming increasingly recognized as an indispensable component of the cloud computing value chain. As the number of cloud service providers grows and the requirements of consumers become more complex, the need for actors to assume the role of an intermediary between providers and consumers will only become bigger. In the future, enterprises will demand brokerage capabilities that will go far beyond what cloud service intermediaries offer today. Continuous quality assurance will be one such type of brokerage capability. Two facets of continuous quality assurance intermediation that we expect to see becoming increasingly important are service governance and quality control, and service failure prevention and recovery. In this paper we attempted to introduce all these concepts and to provide a glimpse of related work that one can build on.

Our future work in cloud service governance will focus on realizing a new approach to policy-based management that overcomes many of the limitations of existing cloud service governance solutions discussed in section 4.1. This new approach will build on Linked Data principles and Semantic Web technologies [47], and will comprise three major components: 1) a platform governance ontology serving as the basic vocabulary for describing platform resources and governance policies; 2) mechanisms to generate semantic descriptions of platform resources by means of transformations from their native representation into Linked Data; 3) a methodology to encode all of the different kinds of cloud service policies in some appropriate logic-based form, based on the same ontology, and a generic and reusable abstract descriptions of service resources are infrastructure to check if the conformant to those policies. The approach sought extends to both lifecycle and artefact governance policy and has the advantage of being formal, declarative and fully standards-based, such that cloud service consumers, providers and intermediaries can automate many policy engineering tasks and easily implement changes in the way software lifecycles and artefacts are governed.

In the context of cloud service failure prevention and recovery, our future work will focus on implementing a cross-layer event-based monitoring engine for measuring many kinds of non-functional and functional properties [48], in order to reactively or proactively detect failures. This may also result in different kinds of agreements that govern a fusion of quality characteristics across the cloud stack [49]. Although little work has been done on the engineering of cross-layer M&A across the infrastructure, platform and application layers of the cloud stack, there exists a plethora of approaches for monitoring and adaptation in the field of self-adaptive service-based systems (SBS) that can be borrowed from. Therefore and in order to allow the future engineering of cross-layer M&A, it will be necessary to implement an architecture that will enable the loosely coupled integration of existing monitoring and adaptation approaches and to ensure the propagation of information between them, as well as an event schema might be proved useful for standardising the information conveyed [46]. Finally, appropriate complex event processing techniques will be utilised as a means to analyse collected events, such that failures are detected either as soon as they occur (reactively), using predefined rules to infer such situations, or near before they occur (proactively), by detecting that a sequence of events (event trace or pattern) will be potentially occur.

Acknowledgments. The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n°328392, the Broker@Cloud project (www.broker-cloud.eu).

References

- Defining Cloud Services Brokerage: Taking Intermediation to the Next Level. Gartner (2010)
- 2. Benoit, J., Lheureux, D., Plummer, C.: Cloud Services Brokerages: The Dawn of the Next Intermediation Age, Gartner (2010)
- 3. Bratanis, K., Kourtesis, D., Paraskakis, I., Verginadis, Y., Mentzas, G., Simons, A. J.H., Friesen, A., Braun, S., Horn, G.: A Research Roadmap for Bringing Continuous Quality

Assurance and Optimization to Enterprise Cloud Service Brokers, eChallenges 2013, Dublin, Ireland (2013)

- Kourtesis, D., Paraskakis, I., Simons, A.J.H.: Policy-driven governance in cloud application platforms: an ontology-based approach, Proc. 4th. Int. Workshop on Ontology-Driven Information Systems Engineering, eds. S. de Cesare, F. Gailly, G. Holland, M. Lycett and C. Partridge, 24-27 July, (2012), (Graz: CEUR Workshop Proceedings)
- 5. Marks, E.A.: Service-Oriented Architecture Governance for the Services Driven Enterprise. Hoboken, New Jersey: John Wiley & Sons (2008)
- Zhang, L., Zhou, Q.: CCOA: Cloud Computing Open Architecture. In: Proc. 2009 IEEE Int. Conf. on Web Services (ICWS 2009), pp.607-616, IEEE (2009)
- 7. Brown, W.A., Moore, G., Tegan, W.: SOA governance—IBM's approach. IBM Corporation, White Paper, (2006)
- 8. MuleSoft Galaxy, http://www.mulesoft.org/galaxy
- 9. WSO2 Governance Registry, http://wso2.com/products/governance-registry/
- 10. IBM WebSphere Service Registry and Repository, http://www.ibm.com/software/integration/wsrr/
- 11. Oracle Service Registry, http://www.oracle.com/us/technologies/soa/service-registry-066462.html
- 12. Kourtesis, D.: Towards an ontology-driven governance framework for cloud application platforms, Dept. Comp. Sci., Univ. Sheffield, UK, Tech. Rep. CS-11-11, (2011)
- Fisler, K., Krishnamurthi, S., Dougherty, D. J.: Embracing policy engineering. In: Proc. of the FSE/SDP Workshop on Future of Software Engineering Research (FoSER '10), New York, pp. 109-110, ACM (2010)
- Uszok, A., Bradshaw, J., Jeffers, R., Johnson, M., Tate, A., Dalton, J., Aitken, S.: KAoS Policy Management for Semantic Web Services. In: IEEE Intelligent Systems, vol. 19, no. 4, pp. 32-41, IEEE (2004)
- Damianou, N., Dulay, N., Lupu, E., Sloman, M.: The Ponder Policy Specification Language. In: Proc. of Int. Workshop on Policies for Distributed Systems and Networks (POLICY 2001), Lecture Notes in Computer Science, vol. 1995, Springer-Verlag, (2001)
- Kagal, L., Finin, T., Johshi, A.: A Policy Language for a Pervasive Computing Environment. In: Proc. of the 4th IEEE Int. Workshop on Policies for Distributed Systems and Networks (POLICY '03), Washington, DC, USA, pp. 63-74, IEEE Computer Society (2003)
- Kolovski, V., Parsia, B., Katz, Y., Hendler, J.: Representing Web Service Policies in OWL-DL. In: Proc. of 4th Int. Semantic Web Conf. (ISWC 2005), Lecture Notes in Computer Science, vol. 3729, pp. 461-475, Springer-Verlag (2005)
- Kolovski, V., Parsia, B.: WS-Policy and beyond: application of OWL defaults to Web service policies. In: Proc. of the 2nd Int. Semantic Web Policy Workshop (SWPW'06), Athens, GA, U.S.A. (2006)
- Gasevic, D., Kaviani, N., Milanovic, M.: Ontologies and Software Engineering. In: Handbook on Ontologies, 2nd ed., Berlin, pp. 593–615, Springer-Verlag (2009)
- 20. Ontology-Driven Apps Using Generic Applications, http://www.mkbergman.com/948/ontology-driven-apps-using-generic-applications/
- 21. M. Uschold, "Ontology-Driven Information Systems: Past, Present and Future," in Proc. of the 5th Int. Conf. on Formal Ontology in Information Systems (FOIS 2008), C.

Eschenbach and M. Gruninger, Eds., Amsterdam, The Netherlands: IOS Press, 2008, pp. 3-18.

- 22. P. Hitzler, M. Krotzsch, S. Rudolph, Foundations of Semantic Web Technologies, Chapman & Hall/CRC, 2009.
- 23. M. Feridun, A. Tanner, Using linked data for systems management, IEEE Network Operations and Management Symposium (NOMS), 2010.
- 24. M. C. Huebscher and J. A. McCann, "A survey of autonomic computing—degrees, models, and applications," ACM Computing Surveys, vol. 40, no. 3, pp. 1-28, Aug. 2008.
- M. Salehie and L. Tahvildari, "Self-adaptive software: Landscape and research challenges," ACM Transactions on Autonomous and Adaptive Systems, vol. 4, pp. 1-42, 2009.
- A. Bertolino, A. Calabro, F. Lonetti, and A. Sabetta, "GLIMPSE: a generic and flexible monitoring infrastructure," in Proceedings of the 13th European Workshop on Dependable Computing, New York, NY, USA, 2011, pp. 73–78.
- 27. L. Romano, D. De Mari, Z. Jerzak, and C. Fetzer, "A Novel Approach to QoS Monitoring in the Cloud," 2011, pp. 45-51.
- E. Di Nitto, C. Ghezzi, A. Metzger, M. Papazoglou, and K. Pohl, "A journey to highly dynamic, self-adaptive service-based applications," Automated Software Engineering, vol. 15, pp. 313-341, 2008.
- 29. D. Dranidis, E. Ramollari, and D. Kourtesis, "Run-time Verification of Behavioural Conformance for Conversational Web Services," in Proceedings of the 7th IEEE European Conference on Web Services (ECOWS), 2009.
- M. Oriol Hilari, J. Marco Gomez, X. Franch, and D. Ameller, "Monitoring Adaptable SOA Systems using SALMon," in 1st Workshop on Monitoring, Adaptation and Beyond (MONA+), 2008, pp. 19-28.
- B. Wetzstein, P. Leitner, F. Rosenberg, S. Dustdar, and F. Leymann, "Identifying influential factors of business process performance using dependency analysis," Enterp. Inf. Syst., vol. 5, no. 1, pp. 79–98, 2011.
- L. Baresi, M. Caporuscio, C. Ghezzi, and S. Guinea, "Model-Driven Management of Services," in 2010 IEEE 8th European Conference on Web Services (ECOWS), 2010, pp. 147-154.
- Kazhamiakin, R., S. Benbernou, L. Baresi, P. Plebani, M. Uhlig and O. Barais. Adaptation of Service-Based Systems. Service Research Challenges and Solutions for the Future Internet, Lecture Notes in Computer Science, 2010, Volume 6500/2010, 117-156 (2010).
- M. Papazoglou, K. Pohl, M. Parkin, and A. Metzger, Eds., Service research challenges and solutions for the future internet: S-cube - towards engineering, managing and adapting service-based systems. Berlin, Heidelberg: Springer-Verlag, 2010.
- Moser, O, Rosenberg, F & Dustdar, S. Non-Intrusive Monitoring and Service Adaptation for WS-BPEL, In Proceedings of the International Conference on World Wide Web, pp 815-824 (2008).
- Baresi, L, Guinea S & Plebani, P. Policies and Aspects for the Supervision of BPEL Processes. Advanced Information Systems Engineering, Lecture Notes in Computer Science, Vol. 4495, pp. 340-354 (2007).
- Siljee, J, Bosloper, I, Nijhuis, J & Hammer, D. DySOA: Making Service Systems Selfadaptive, In the Proceedings of the 3rd International Conference Service-Oriented Computing – ICSOC, pp. 255–268 (2005).

- Gehlert, A, Metzger, A, Karastoyanova, D, Kazhamiakin, R, Pohl, K, Leymann, F & Pistore, M. Chapter 6: Integrating Perfective and Corrective Adaptation of Service-based Applications, In Service Engineering: European Research Results Book by Dustdar S & Li F, Springer, pp. 137-169 (2011).
- Hielscher, J, Kazhamiakin, R, Metzger, A & Pistore, M. A framework for proactive selfadaptation of service-based applications based on online testing, In the Proceedings of the ServiceWave 2008, pp. 122-133 (2008).
- D. Dranidis, A. Metzger, and D. Kourtesis, "Enabling Proactive Adaptation through Justin-Time Testing of Conversational Services," in Towards a Service-Based Internet, vol. 6481, E. Nitto and R. Yahyapour, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 63-75.
- 41. Charfi, A & Mezini, M.Ao4bpel: An aspect-oriented extension to bpel', World Wide Web, vol. 10, no. 3, pp. 309–344 (2007).
- 42. Charfi, A, Dinkelaker, T. & Mezini, M. A plug-in architecture for self-adaptive web service compositions, Proceedings of the 2009 IEEE International Conference on Web Services (ICWS '09), IEEE Computer Society, pp. 35–42 (2009).
- Karastoyanova, D & Leymann, F. BPEL'n'Aspects: Adapting Service Orchestration Logic, Proceedings of the IEEE International Conference on Web Services, ICWS 2009, pp. 222 – 229 (2009).
- 44. Hermosillo, G, Seinturier L & Duchien L. Using Complex Event Processing for Dynamic Business Process Adaptation, In Proceedings of the IEEE SCC 2010, pp 466-473 (2010).
- 45. A. Marconi, A. Bucchiarone, K. Bratanis, A. Brogi, J. Camara, D. Dranidis, H. Giese, R. Kazhamiakin, R. Lemos, C. Marquezan, and A. Metzger, "Research challenges on multi-layer and mixed-initiative monitoring and adaptation for service-based systems," in 2012 Workshop on European Software Services and Systems Research Results and Challenges (S-Cube), ICSE 2012.
- 46. Bratanis, K., "Towards engineering multi-layer monitoring and adaptation of servicebased applications", Dept. Comp. Sci., Univ. Sheffield, UK, Tech. Rep. CS-12-04, 2012.
- 47. Heath, T., and Bizer, C., (2011). Linked Data: Evolving the Web into a Global Data Space. In Synthesis Lectures on the Semantic Web: Theory and Technology, vol. 1, no. 1, Morgan and Claypool, 2011, pp. 1-136.
- Bratanis, K., Dranidis, D., Simons, A.J.H.: An Extensible Architecture for Run-time Monitoring of Conversational Web Services. Proceedings of the 3rd International Workshop on Monitoring, Adaptation and Beyond / ECOWS 2010. pp. 9–16. ACM (2010).
- Bratanis, K., Dranidis, D., Simons, A.J.H.: SLAs for cross-layer adaptation and monitoring of service-based applications: A case study. Proceedings of the International Workshop on Quality Assurance for Service-Based Applications. p. 28. ACM Press (2011).

Cloud Service Portability Opportunities

Magdalena Kostoska, Marjan Gusev, Sasko Ristov, and Kiril Kjiroski

Ss. Cyril and Methodius University Faculty of Information Sciences and Computer Engineering Rugjer Boshkovik 16, PO Box 393, Skopje, Macedonia {magdalena.kostoska,marjan.gushev, sashko.ristov,kiril.kjiroski}@finki.ukim.mk

Abstract. Several open research questions motivated us to conduct this research: How to wisely choose a cloud host for services and how to change the cloud provider in an easy manner. Service portability in the cloud is the ability to move the services from one to another provider in a simpler manner. It is still an open research area which deserves more attention.

This paper gives an overview of the current opportunities in the area of cloud service portability, while giving special attention to TOSCA's approach. Our research is based on an *adapter model* to achieve the cloud portability. It describes the foundation of a framework which will allow the services to be described with the usage of XML. The solution is to build a custom platform (PaaS) capable of accepting installations of any portable application. The application provider will have to describe: 1) the installation requirements and application structure in separate XML files and 2) the platform to perform the necessary installation according to the XML files, and 3) to automatically deploy the application/services.

Keywords: Cloud computing interoperability, Cloud service portability, TOSCA

1 Introduction

Although there are currently a lot of available cloud providers, we expect their number to grow immensely in the future. It is very important to make the right decision when choosing the cloud platform, since migrating services to another platform is not an easy task. The open interoperability and portability issues not only concern the services, but also the data these services have already collected and processed.

There are no published or fully developed solutions for these problems. The problem is exposed even more due to the lack of large-scale adopted cloud standards. In this situation, while waiting for the emergence of real standards, each vendor tries to impose newly developed technical solutions as a de-facto standard. The portability of cloud services in each of the layers of the service stack is discussed in a lot of papers from the perspective of standards [6, 18] and has even been addressed in open standards [4]. Several solutions that have been proposed include the abstraction-driven approach [14] and CSAL [5], while other approaches focus on exploiting the semantic technology [7, 2, 13, 8].

The efforts to construct cloud standards in this area are represented by the initiative organized by the OASIS (Organization for the Advancement of Structured Information Standards) task force TOSCA (Topology and Orchestration Specification for Cloud Applications) technical committee, that works on the standard for "interoperable description of application and infrastructure cloud services, the relationships between parts of the service, and the operational behavior of these services (e.g., deploy, patch, shutdown)" [10]. Although some influential companies like IBM, SAP, Citrix and RedHat are part of this process there is no guarantee that this standard will be accepted by the other vendors [1].

In this paper we present and analyze the TOSCA model and propose a new platform to enhance service portability.

2 Topology and Orchestration Specification for Cloud Applications (TOSCA)

OASIS as a not-for-profit, international consortium, aims to develop, converge and adopt open standards for the global information society. The OASIS TOSCA Technical Committee works to enhance the portability of cloud applications and services. Their goal is to enable the interoperable description of application and infrastructure cloud services, the relationships between parts of the service, and the operational behavior of these services [10].

The open standards are used to describe a service and how to manage it independently from the supplier who creates the service , as well as from any particular cloud provider (or its technology) that hosts the service. This technical committee has produced a specification (Version 1.0) which explains how to define services in a portable and interoperable manner by a Service Template document. The TOSCA language introduces a grammar for describing service templates by means of Topology Templates and plans [10]. TOSCA utilizes XML Schema 1.0⁻¹ and WSDL 1.1 [16]. It also contains non-normative references to BPEL 2.0 (Web Services Business Process Execution Language Version 2.0, OASIS Standard) [9], BPMN 2.0 (OMG Business Process Model and Notation Version 2.0) [12], OVF (Open Virtualization Format Specification Version 1.1.0) [3] and XPATH 1.0 (XML Path Language Version 1.0, W3C Recommendation) [17].

The TOSCA specification uses TOSCA xml [11] and xs [15] namespace prefixes, but allows extensibility from other namespaces for attributes and elements,

 $^{^1}$ Defined by W3C: XML schema part 1 and 2, Name spaces in XML 1.0 and XML base

which do not oppose the semantics of the TOSCA namespace [10]. The specification defines a meta model for defining the structure of an IT service and its management. The Topology Template defines the structure of a service. Plans define the process models that are used to create, terminate and manage a service. The major elements defining a service are depicted in Figure 1.

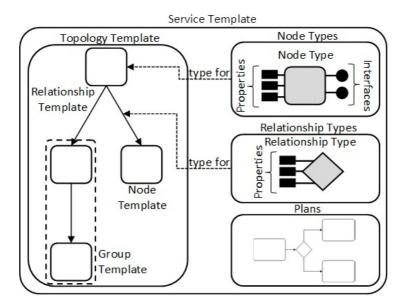


Fig. 1. TOSCA Service Template [10]

The specification supports the following major use cases:

- Services as Marketable Entities,
- Portability of Service Templates,
- Service Composition and
- Relation to Virtual Images.

3 A New Proposed Model

This model is based on a specific platform which serves as a PaaS. It uses the universal XML language to describe the installation requirements as well as the application/services structure. Its goal is to provide an easy migration of SaaS applications and services from one provider/solution to another via a specific custom platform.

3.1 Platform

This platform is intended to be deployed on either a public or a private cloud and can be deployed using multiple hypervisors. The underlying operating system is planned to be Linux. In the first phase we plan this platform to be able only to host multiple applications and provide data and messaging services. It will offer a limited list of web servers, languages, data and messaging services. Our target will be the open-source and freeware solutions in the first experimentation phase. Later we plan to enhance this platform with appropriate tools to control the deployed application and to monitor the work of these applications. We are also considering the possibility of separating different groups of applications in dedicated allocations.

It will use the universal XML language to describe the installation requirements as well as the application/services structure and we will develop and use dedicated tools for processing of these files. These tools will assist with the installation of the required tools and software and with the deployments of the specified application/services.

3.2 XML description of installation requirements

The list of software that could be installed will be predefined and the user responsibility will consist of defining XML file of the required software, chosen from the offered list. A XML file is created according to the user's choice.

We recommend using open-source or freeware tools in this step in order to automate the installation. An example of such a tool is IzPack, which is a Java based software installer builder that will run on any operating system that contains a Java Virtual Machine (JVM). It is a software covered by the GNU General Public License (GPL) and can be modified.

The root element of an XML installation file is *<*installation> and some more important tags that can be used are:

- <info> This element is used to specify general information such as application name and version, website etc. for the installer.
- <variables> and <dynamicvariables> This elements allow the definition of variables and dynamic variables for the variables substitution system.
- <conditions > - This element allows the definition of conditions which can be used to dynamically change the installer
- <resources > - This tag contains paths to resources used by various parts of the installer

This tool also offers the creation of a panel, so the user can be aware of the installation progress.

3.3 XML description of applications/services structure

To describe applications/services structure with XML, we plan to use the Service Template defined by TOSCA. This template represents a deployment Topology Template. It consists of Node Templates and Relationship Templates that together define the topology model of a service as a directed graph. We will not use the Plans, which describe the management aspects of service instances.

The service template document has a lot of properties, although most of them are optional. A property that has to be included in each template is *id*, an identifier which will be unique and TopologyTemplate which defines the structure (the relation between those components) of the application. The TopologyTemplate consists of a NodeTemplate which specifies the components and a RelationshipTemplate which specifies the relationships between the components. It can also define tags for description and boundary definitions like properties and their constraints, requirements, capabilities, policies and interfaces. The Node-Template consists of at least the following obligational properties: id; type, a QName value; and properties and their constraints. It may also define requirements, capabilities, policies and deployment artifacts. The RelationshipTemplate must define id and type, as well as source and target elements. It may also describe the relationship constraint type.

The following code represents a part of a service template defined by TOSCA. Example of a part of a Service Template given by TOSCA

```
<ServiceTemplate id="MyService"</pre>
               name="My Service">
 <TopologyTemplate>
    <NodeTemplate id="MyApplication"
                  name="My Application"
                  type="my:Application">
      <Properties>
        <ApplicationProperties>
          <Owner>Frank</Owner>
          <InstanceName>Thomas favorite application</InstanceName>
        </ApplicationProperties>
       </Properties>
    </NodeTemplate>
    <NodeTemplate id="MyAppServer"
                  name="My Application Server"
                  type="my:ApplicationServer"
                  minInstances="0"
                  maxInstances="unbounded"/>
    <RelationshipTemplate id="MyDeploymentRelationship"
                          type="my:deployedOn">
      <SourceElement ref="MyApplication"/>
      <TargetElement ref="MyAppServer"/>
    </RelationshipTemplate>
```

```
</TopologyTemplate>
```

</ServiceTemplate>

3.4 Usage

The customer provides the requirements for the applications/services, while the required software is automatically installed. Two different scenarios follow:

- 1. the customer gives an XML document that describes the application, the services, the relationships between the services and the source code, so the application is automatically deployed;
- 2. the customer develops the application on our platform and in parallel with this activity produces the required XML file. Whenever the customer wants to change the cloud provider a new instance of the platform is deployed on the new cloud solution and two scenarios are possible.

Scenario 1 (Relies on direct communication between the old and the new instance): The customer enters the location of the previous (still active) instance on the new instance and pre-installed software, so the migration starts automatically. Firstly, the new instance of the platform takes the defined specifications and installs the necessary software. Afterwards, it reads the application/services structure file and it automatically deploys this solution on the new instance.

Scenario 2 (Relies on communication between the customer and the new instance, using data from the old instance): First, on the old instance, the customer performs export of the current state of the platform and the package of data is downloaded to the customer's local machine. Then the customer connects to the new instance, runs the preinstalled software and uploads the data, which the software will process and will also perform installation and deployment.

4 Future Research Directions and Work

In the first phase we will use OpenStack and OpenNebula cloud solutions to set the first instances of the described platform and we will try to deploy services on one instance and then migrate these services to other instances. The services will be deployed on Apache Tomcat web server and will be built with usage of the Java programing language. In this phase we also plan to use MySQL and PostgreSQL as data services.

In the second phase we plan on adding more types of web servers, while including more languages and improving the platform with monitor tools. Furthermore, we plan on providing the separation of different groups of applications, as well as providing automatic deployment of multiple instances of the same application.

In the future we plan to develop a model for data migration and to enable proprietary software usage.

5 Conclusion

Cloud service portability is not yet provided in a standardized manner. TOSCA represents a first step in this direction. This specification should standardize the portable deployment and the migration of existing applications onto the cloud. Unfortunately it does not represent a standard, because it is still in the process of development, nor does it guarantee acceptance by the major cloud providers like Google and Amazon, which are not taking part in the creation of this specification.

Our solution, on the other hand, will represent an adapter solution in a manner that it is not intended to be standard, but transitional solution until that kind of standard arises (if ever). It uses part of the TOSCA specification, but is yet independent of this specification. Even if SaaS cloud standardization ever becomes possible, this solution will have already offered a XML description of the applications, thus having eased the transformation of this description into the new standard format. This solution promotes service portability and in that way it also promotes service interoperability in the cloud. It will also ease the migration of the services and will be provider independent. It uses PaaS as a base since at this moment it is not possible to enable smooth migration on SaaS level due to the lack of standardization on any level of the cloud computing service stack.

References

- Binz, T., Breiter, G., Leyman, F., Spatzier, T.: Portable cloud services using TOSCA. Internet Computing, IEEE 16(3), 80–85 (2012)
- Cretella, G., Di Martino, B.: Towards automatic analysis of cloud vendors APIs for supporting cloud application portability. In: Sixth International Conference on Complex, Intelligent and Software Intensive Systems (CISIS). pp. 61 – 67 (2012)
- 3. DFTM: Open virtualization format specification version 1.1.0 (Jan 2010), http://www.dmtf.org/sites/default/files/standards/documents/ DSP0243_1.1.0.pdf
- 4. Galan, F., Sampaio, A., Rodero-Merino, L., Loy, I., Gil, V., Vaquero, L.M.: Service specification in cloud environments based on extensions to open standards. In: Proceedings of the Fourth International ICST Conference on COMmunication System softWAre and middlewaRE. ACM New York (2009)
- Hill, Z., Humphrey, M.: CSAL: A cloud storage abstraction layer to enable portable cloud applications. In: IEEE Second International Conference on Cloud Computing Technology and Science (CloudCom). pp. 504 – 511 (2010)
- Lewis, G.: Role of standards in cloud-computing interoperability. In: 46th Hawaii International Conference on System Sciences (HICSS). pp. 1652 – 1661 (2013)

- Loutas, N., Kamateri, E., Tarabanis, K.: A semantic interoperability framework for cloud platform as a service. In: IEEE Third International Conference on Cloud Computing Technology and Science (CloudCom). pp. 280 – 287 (2011)
- Moscato, F., Aversa, R., Di Martino, B., Fortis, T.: An analysis of mO-SAIC ontology for cloud resources annotation. In: Federated Conference on Computer Science and Information Systems (FedCSIS). pp. 973 – 980 (2011)
- 9. OASIS: Web services business process execution language version 2.0 (Apr 2007), http://docs.oasis-open.org/wsbpel/2.0/wsbpel-v2.0.html
- OASIS: Topology and orchestration specification for cloud applications version 1.0. (Mar 2013), http://docs.oasis-open.org/tosca/TOSCA/v1.0/ cs01/TOSCA-v1.0-cs01.html
- 11. OASIS: XML Namespace Document for TOSCA Version 1.0 (Mar 2013), http://docs.oasis-open.org/tosca/ns/2011/12
- OMG: Business Process Model and Notation (BPMN) Version 2.0 (Jan 2011), http://www.omg.org/spec/BPMN/2.0/
- 13. Petcu, D.: How to build a reliable mOSAIC of multiple cloud services. In: Proceedings of the 1st European Workshop on Dependable Cloud Computing (2012)
- Ranabahu, A., Maximilien, E., Sheth, A., Thirunarayan, K.: Application portability in cloud computing: An abstraction driven perspective. IEEE Transactions on Services Computing PP(99), 1 (2013)
- W3 consortium: XML Schema (Mar 2013), http://www.w3.org/2001/ XMLSchema
- 16. W3C: Web Services Description Language (WSDL) 1.1 (Mar 2011), http: //www.w3.org/TR/wsdl
- 17. WC3: XML Path Language (XPath) Version 1.0 (Nov 1999), http://www. w3.org/TR/1999/REC-xpath-19991116/
- Zhang, Z., Wu, C., Cheung, D.W.: A survey on cloud interoperability: taxonomies, standards, and practice. ACM SIGMETRICS Performance Evaluation Review 40(4), 13–22 (2013)

Performing a Penetration Test in a Data Network

Renalda Kushe¹, Adrian Shehu¹

¹ Faculty of Information Technology. Sheshi Nene Tereza, Nr. 4, Tirane, Albania <u>rkushe@fti.edu.al, adshehu@tcn.al</u>

Abstract. Nowadays the information is very valuable and there is increasing dependence on ubiquitous access to it. This need for information makes it vulnerable to various threats and attacks. A penetration test is a way of detecting vulnerabilities that exist in a system or network in order to increase upper management awareness of security issues and either to test intrusion detection and response capabilities. In this paper, we will present a penetration test performed in a data network. The aim of this paper is to identify and to overcome the vulnerabilities of a data network.

Keywords: Penetration test, attack, vulnerability, exploit, defense, security.

1 Introduction

The Internet has made a big impact in our lives and society in general. It will continue to grow while people find more uses for it every day. As the Internet keeps growing, problems continue growing as well. One of these "problems" is the security issue. Personal information is shared across numerous distributed sites on the web. But how protected this information is? The privacy issue is a big concern as anyone can just see your details e.g. your age, address and where you live, e-mail address etc. The danger is ones someone get hold of your details they can be easily used for anything e.g. criminal activities, fraud etc. The same effect could be seen in computer systems used for global commerce. Altering credit card transaction data at e-commerce Websites could cause the altered information to spread into banking systems, thus eroding public confidence in the financial sector.

Nowadays the information is very valuable and there is increasing dependence on ubiquitous access to it. This need for information makes it vulnerable to various threats and attacks.

A penetration test (PenTest) [1] is a way of detecting vulnerabilities that exist in a system or network in order to increase upper management awareness of security issues and either to test intrusion detection and response capabilities too.

There are a variety of reasons for performing a penetration test. One of the main reasons is to find vulnerabilities and fix them before an attacker does.

The penetration tests enter to form part of the securing processes, maintenance and evaluation of information system. Their mission consists of proving the implanted policies of security in a technological platform, using a assembly of techniques that result in an analysis that demonstrates all the weaknesses that put in risk the confidentiality, integrity and availability of the information, giving a start point to proceed to the accomplishment of adjustments in the security policies and generating a feedback with other processes, the real knowledge of the present situation and determining the optimization and/or application of new technologies and processes that increase the reliability degree.

There are two types of penetration test: External and Internal PenTests. These will be described in the following sessions respectively 1.1 and 1.2. In paragraph 2 will be discussed the PenTest phases; Planning, Exploitations, Reporting respectively in sessions 2.1, 2.2 and 2.3. The Experimental Setup is described in paragraph 3 and in paragraph 4 some conclusions of this work are given.

1.1 External PenTests

An external PenTest is performed from the outside of the data network that will be tested. This data network is reachable for the tester through the internet connections. The data network traditionally is protected with a firewall.

The testing is focused on all components of target system (TOS) including servers, infrastructure and the underlying software comprising the target. It may be performed with no prior knowledge of the site (method known as black box) or with full disclosure of the topology and environment (method known as white box). This type of testing consists of comprehensive analysis of publicly available information about the target, a network enumeration phase where target hosts are identified and analyzed and the behavior of security devices such as screening routers and firewalls are analyzed. Vulnerabilities and mis-configurations within the target hosts should then be identified, verified and the implications assessed.

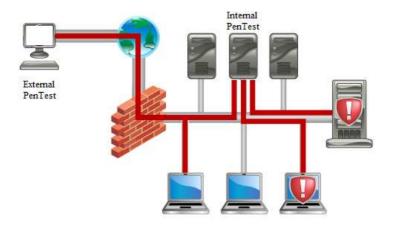


Fig. 1. Two types of PenTest (internal and external PenTests)

1.2 Internal PenTests

An internal PenTest is performed from the inside of the data network that will be tested. In this case, the tester is one of the nodes in the network.

The internal PenTest follows a similar methodology to external testing, but provides a more complete overview of the overall security. Testing is typically performed from a number of network access points, representing each logical and physical segment.

2 PenTest Phases

There must be a clear objective for the penetration test to be conducted. In most cases the objective of a penetration test is to demonstrate that exploitable vulnerabilities exist within an organization's network infrastructure. The scoping of the penetration test is done by identifying the machines, systems and network, operational requirements and the staff involved.

A PenTest must be performed through the following phases: Planning, Exploitation, Reporting. These phases will be discussed in detail in the following sessions, respectively is session 2.1, 2.2 and 2.3.



Fig. 2. PenTest Phases (three phases: planning, exploitation, reporting)

These phases should be performed in the linear order (which mean first the Planning phase, then Exploitation and last Reporting), however they interconnect with each other. The Fig. 3 shows how these phases interconnect. As the Fig.3 shows, the

Reporting phase interconnects with Exploitation phase also with Planning phase too; this is because the planning phase conclusions should be reported.

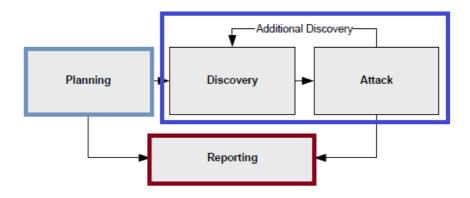


Fig. 3. The interconnection between phases

2.1 Phase 1: Planning

The first step on performing a PenTest is to gather as much information as possible about the targeted systems, networks and people too. There's a wealth of tools and online resources available for us to do the necessary information gathering. If the intended target has an online website, this is a good place to start your information gathering. We should always remember that any kind of information gathered during this stage may prove useful to us in the other stages of the penetration test. Information is power.

The Planning phase should follow these sub-phases:

- Information gathering
- Detecting live systems
- Scanning

Gathering the necessary information is realized in two aspects:

- Technical information (systems and networks)
- Social engineering information

The social engineering information is performed because not always the attacks are technical. Social Engineering addresses intrusions without specialized technical abilities. It relies on human interaction and involves tricking other people into breaking normal security procedures. Social engineering usually involves a scam; trying to gain the confidence of a trusted source by relying on the natural helpfulness of people as well as their weaknesses, appealing to their vanity, their authority and eaves dropping are natural techniques used. Other techniques involve searching refuse bins for valuable information, memorizing access codes by looking over someone's shoulder, or taking advantage of people's natural inclination to choose passwords that are meaningful to them but can be easily guessed.

Detecting live systems is about to find the number of systems that are reachable. The expected results that should be obtained from a network surveying should consist of domain names, server names, Internet service provider information, IP addresses of hosts involved as well as a network map. A network survey will also help us to determine the domain registry information for the servers. This allows us to check and see the range of IP addresses that are owned by the targeted organization.

Upon doing a network survey and obtaining adequate information about the network the next task to be done would be to do a port scanning to obtain information about closed and open ports running on the systems or network.

2.2 Phase 2: Exploitation

The phase 2 should follow these sub- phases:

- Vulnerability research
- Exploitation

After having gathered the relevant information about the targeted system, the next step is to determine the vulnerability that exists in each system. Penetration testers should have a collection of exploits and vulnerabilities at their disposal for this purpose. The knowledge of the penetration tester in this case would be put to test. An analysis will be done on the information obtained to determine any possible vulnerability that might exist. This is called manual vulnerability scanning as the detection of vulnerabilities is done manually.

After determining the vulnerabilities that exist in the systems, the next stage is to identify suitable targets for a penetration attempt. The time and effort that need to put in for the systems that have vulnerabilities need to be estimated accordingly.

2.3 Phase 3: Reporting

After conduction all the tasks above, the next task ahead is to generate a report. The report should start with an overview of the penetration testing process done. This should be followed by an analysis and commentary on critical vulnerabilities that exist in the network or systems. Vital vulnerabilities are addressed first to highlight it to the organization. Less vital vulnerabilities should then be highlighted.

The phase 3 should follow these sub- phases:

- Finding Analyses
- Risk Rating
- Reporting

The report is a document that compiles all the information obtained in the different phases, summarizes the analysis of the results and the documentation of the tests, in addition it raises the recommendations that contribute to the cycle of continuous improvement, contributing to the adjustment in the policies of established security and consequently to the implanted security system, this with the purpose of controlling those services that were not considered initially or which they are totally new within the scheme and which they represent potentially alternative of danger against integrity, confidentiality and availability of the information.

3 Experiment Setup

This security experiment setup covers a penetration test performed over a data network. This assessment is carried out from an internal PenTest perspective (described in paragraph 1.2) with the only supplied information being the tested hosts IP addresses range. No other information was assumed at the start of the assessment.

In the phase1 (described in paragraph 2.1), we detect the live systems, their O.S and determine the running services and their versions. To perform phase 1, is utilized the Nmap software [2]. Fig. 4 shows the Nmap scanning results.

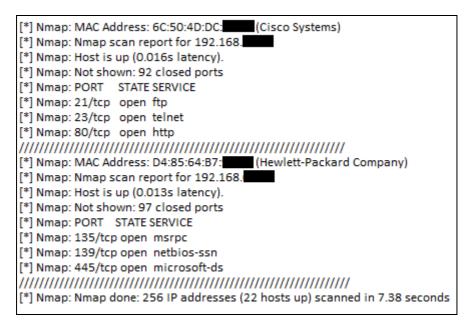


Fig.4 Nmap Scanning

In the phase 2 (described in paragraph 2.2) we utilize the information gathered in phase1. We start to find the vulnerability for each O.S and service that we discovered and after that trying to exploit it. In the figure 5 are presented all the vulnerabilities for this network, however not all the founded exploits are executable.

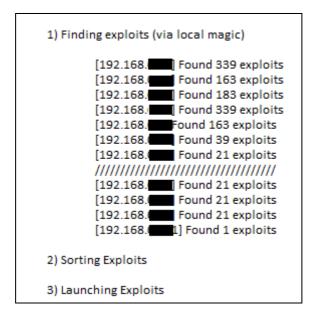


Fig. 5 Finding vulnerabilities

In this phase we will execute an exploit code but first we have to scan for finding the executable exploits. To perform this operation, we will use one of the above scanned hosts with 21 vulnerabilities. The targeted system to exploit is the HP PC shown in Fig. 4. To perform phase 2, is utilized the Metasploit Framework [3][5][6].

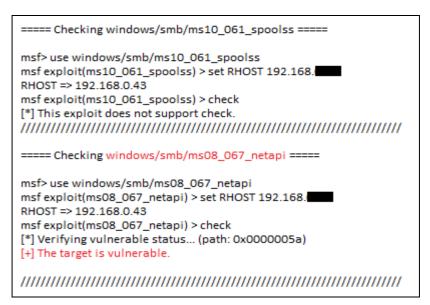


Fig 6 Finding for executable exploits

So, as we can see from the Figure, the exploit windows/smb/ms08_067_netapi is executable.

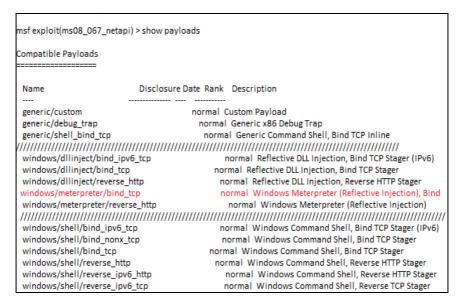


Fig. 7 The exploit code used



Fig. 8 After the exploitation

The exploit results for this exploit code are privilege escalation. By implementing this exploit, the user receives privileges he is not entitled to. These privileges can be used to delete files, view private information, or install unwanted programs such as viruses etc. The rank for this exploit code is "Good". [4]

In phase 3 (described in paragraph 2.3) we have to report the level of security and if is required to perform defense techniques to improve the security level. The Fig. shows an improvement of numbers of threads for the exploited system from 21 to 16.

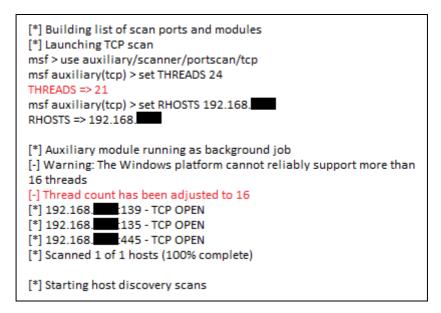


Fig. 9 Performing defense techniques

Based on the results of the first two steps, we start analyzing the results. The analyzed results should be shown in detail in the Report Form. (The Report Form is a document which contains the PenTest results). The main analyze results are the risk rating and the improvement of the security. The risk rating is based on this calculation: Risk = Threat * Vulnerability * Impact [7]

Threat		Lov	v			Medium			High			Critical					
Vulner	ability	L	м	н	С	L	м	н	С	L	М	н	С	L	м	H	С
Impact	Low	1	2	3	4	1	4	6	8	3	6	9	12	4	8	12	16
	Medium	2	4	6	8	4	8	12	16	6	12	18	24	8	16	24	32
	High	3	6	9	12	6	12	18	24	9	18	27	36	12	24	36	48
	Critical	4	8	12	16	8	16	24	32	12	24	36	48	16	32	48	64

Fig. 10 Results Analyze [7]

Based on the above risk rating calculation, in this data network, the thread count from 21 has been adjusted to 16; therefore the threat level has been decreased from medium to low, so it has been an improvement of the security level for this data network.

L	Low	1	16
М	Medium	17	32
н	High	33	48
C	Critical	49	64

Fig. 11 Rating Calculation [7]

4 Conclusions

This paper gives information about the penetration testing, its methodologies and its application. In this paper we have performed a penetration test in a real data network.

There are many security problems for which penetration tests will not be able to identify. This is because of the lack of information gathered. The complete information of systems, network and people often is missing. A penetration test can only identify those problems that it is designed to look for. If a service is not tested then there will be no information about its security or insecurity.

However the PenTest is a powerful defense technique to detect the network vulnerabilities and to improve the security level.

References

- 1. Zina A.: Penetration Testing Techniques for analysis perspective, <u>http://www.scribd.com/doc/13470161</u>, Final-Project-Penetration-Testing.
- 2. <u>http://nmap.org/</u> Nmap scanning tool
- 3. http://www.metasploit.com/ Metasploit Framework.
- 4. <u>http://www.metasploit.com/modules/exploit/windows/smb/ms03_049_netapi</u>, The exploit Rank.
- 5. Signh A.: Metasploit Penetration testing Cookbook. PACKT Publishing, Birmingham-Mumbai (2012)
- 6. Kennedy D., O'Gorman J., Kearns D., Aharoni M.: Metasploit, The penetration Tester's Guide. William Pollock, San Francisco (2011)
- 7. SANS Institute InfoSec Reading Room: Writing a Penetration Testing Report

Green Cloud Computing Trends: A Review

Enida Sheme¹, Neki Frasheri²

^{1,2} Department of Computer Engineering, Polytechnic University of Tirana, Albania ¹ esheme@fti.edu.al, ² nfrasheri@fti.edu.al

Abstract. The recent growth of Cloud Computing technology has received a major attention on the research arena worldwide, mainly being focused on performance issues of greater and larger Data Centers. It is our interest to know about their cost regarding energy and power consumption. Are we getting to the point where it's too late to sustain a green environment, because of the high level of risk we are taking without analyzing? This study tends to give answers to these questions by presenting the current state of the art on green cloud computing. This paper provides an overview of researches and reports published in the latest three years, 2010-2013, in the fields of Cloud Computing Energy Efficiency improvements and awareness. We highlight the trend of researches on this area from the "Cloud Service Providers" and Environmentalists point of view.

Keywords: Green, Energy Efficiency, Power Consumption, Cloud Computing

1 Introduction

In this paper we present an overview of the current research regarding energy efficiency and keeping Cloud Computing as green as it is possible. We want to know if performance is the only thing that matters, what do Cloud Service Providers think and what studies and reports say about this. Our main aim is to give a clue on the importance this area is being given to by researchers, the research trend in the latest years, the impact that awareness reports of the actual energy state are bringing on intensifying new techniques for saving energy in actual Cloud infrastructure, or new designs with energy efficiency in mind. Although there are some similar reviews on this area such as [7], [27], [29], they are mainly focused on causes and problems of high energy consumption and the need for green IT in the Cloud, rather than detecting and highlighting the tremendous growing trend of studies on the field as we do. Also they gather information from studies belonging to years 2005'-2011', meanwhile all our references are published during 2010'-2013' period. We are going to classify all publications gathered for this study according to a methodology explained in section 2, to continue then with results and their analysis.

The rest of this paper is organized as follows: In the second section we present the methodology used for this study, including searching tools and classification scheme

for better examining our publications sample. In the third section we present results and analysis of these searches in a quantitative perspective. Then the fourth section presents results on specific research topic perspective. We conclude finally with the conclusions and discussions we have come preparing and accomplishing this study.

2 Methodology

In order to better assess the state of cloud computing energy (green) research and awareness, we portray a current landscape of these two components: online computing database searching for each of the years 2010-2013 and Conferences on Cloud Computing search. We'll focus on publications and the role of each of the components as follows:

1. First we chose these databases for our search: IEEE search, ACM search, Google search, as most popular online databases for computing publications. We examined papers and reviews and their references too, to result in a total of 40 study samples, papers and reports published during the period 2010 - 2013.

The searching keywords were: energy / power / green cloud, and only after reading the full title of the publication, it was added in the list of our study sample if it was right in the interested area.

2. Secondly we chose four important conferences on Cloud Computing: International Conference on Cloud Computing, IEEE GreenCom Conference, IEEE Cloud, International Conference on Cloud and Green Computing, to come up with the total number of publications per year focusing on our interested area.

We chose a classification framework to categorize our study sample of 40 articles, as shown in Table 1, based on the literature review, the interest of energy and green cloud computing research and the existing ACM Computing Classification System [1]. As we can see from the table 1, after a full text review of all selected publications, we grouped them in two main domains: research papers and reports. Then in these sub domains: Review, Improvements in existing Technology, New Design, regarding first domain, and Environmental Impact, Cloud Service Provider Perspective regarding the second domain.

Domain	Subdomain
1.Research Paper	1) Review
_	2) Improvements in existing technology
	3) New designs
2.Report	1) Environmental Impact
·	2) Cloud Service Provider Perspective

Table 1. Main classification scheme of publications in two domains and five sub domains.

3 Literature review: quantitative results and analysis

In this section we give the results of our study, based on the methodology explained in section 2, so there are two results subsections: IEEE / ACM / Google Search and Cloud Computing Conferences Search.

3.1 IEEE / ACM / Google Search

In this subsection, we will classify all 40 selected publications of the first search in two main categories: Papers and Reports and give the results of our study.

Table 2. Representation of the total number of publications per year regarding energy in cloud computing. This total is then classified in papers vs. reports, and percentage of them representing papers and reports

Year	Total nr. of	Papers / Reports	Percentage Papers /
	publications		Reports(%)
2010	6	4 / 2	67 / 33
2011	10	6 / 4	60 / 40
2012	17	15/2	89 / 11
2013	7	6 / 1	86 / 14
Total	40	31/9	77.5 / 22.5

As we can see from table 2, there has been a continuous growth of publications on the field from the year 2010 by approximately 50% for each year, resulting in a total number almost four times greater by 2013', which we can see in Figure 1.

Also we can identify that 77.5% of all selected publications are research papers and 22.5% of them are reports. We can clearly see from Figure 2 below that till 2011' there has been a constant growth in number of publications in both research papers and reports, meanwhile in 2012' we can notice a descent on the number of reports but 2.5 times greater number of papers publication. We can justify this as an awareness impact of reports towards researchers of the field.

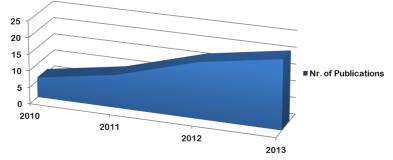


Fig. 1. Total number of publications per year based on Table 2 data. The results for year 2013 are a prediction for all the year, by multiplying the actual April data by 3

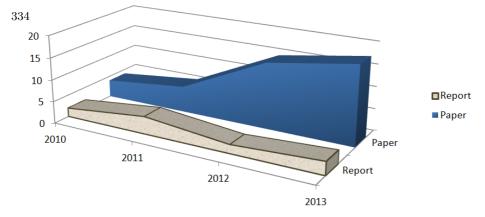


Fig. 2. Total number of reports and papers per year, based on Table 2 data. The results for year 2013 are a prediction for all the year, by multiplying the actual April data by 3.

Based on table 1 classification scheme, we give the results on the total number of publications per sub domain in Table 3 and their percentage on the domain they are part of, shown in Figure 3.

Domain	Sub domain	Nr. of publications out of total	
Papers	Review	5 / 31	
Î	Improvements in existing technology	10 / 31	
	New designs	16 / 31	
Reports	Environmental Impact	6 / 9	
-	Cloud Service Provider Perspective	3 / 9	

Table 3. Representation of the total number of publications per sub domain

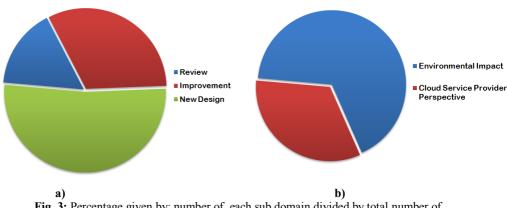


Fig. 3: Percentage given by: number of each sub domain divided by total number of publications per domain, a) Paper domain b) Review domain

3.2 Cloud Computing Conferences Search

In this subsection we present the data we had from four international important Conferences on Cloud Computing with their total number of publications per year regarding energy or green Cloud.

Table 4. Representation of the total number of presented papers per year, regarding directly to energy efficiency and green cloud computing, period 2010-2012

Conference	Nr. of publicat.	Nr.of publicat.	Nr.of publicat.
	2010	2011	2012
International	2	6	14
Conference on			
Cloud Computing			
IEEE GreenCom	-	7	6
Conference			
IEEE Cloud	0	8	12
Conference			
International	-	15	8
Conference on			
Cloud and Green			
Computing			
Total per year	2	36	40

We can see the exploding growth of energy issue on cloud computing in the total number of publications for these four conferences starting from 2 in 2010' to 36 in 2011' and later to 40 in 2012', given in graphical form in Figure 4.

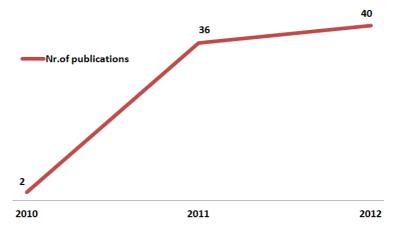


Fig. 4: Total number of publications in four Conferences per year 2010-2012

We have selected this period because the number of publications regarding the energy issue, in 2009' or prior specifically for cloud computing, was almost trivial. Most of 2013 Conferences are still in the "Call for Papers" process along the time this study is made, as such we cannot include information for this year.

4 Literature review: topics-related analysis

This section is an overview of each of the sub domains, showed on table 1, regarding their specific topic on how they contribute to the energy issue. We could not have all full text publications of the four conferences to fully review them, so this analysis is focused on the 40 publications we presented on section 3.1.

For every sub domain we will mention the main strategies, policies, focuses on their research, as well as references and possible analysis.

1. *Research Papers:* This domain includes latest academic researches on the field of energy and power consumption in cloud computing, focusing on new techniques, concepts, designs, or improvements on this area. We categorize them as follows:

1.1 *Review:* this sub domain consist of studies that tend to give a state of the art on energy issue on cloud computing and classifies the existing techniques or hardware and software current solutions on the field, like articles in [7],[26], [28], [29], [33].

1.2 *Improvements in existing Technology:* they tend to focus on the implementation of optimization techniques such as virtualization and consolidation [2], [4], [23], [32], on evaluating cloud computing data centers aspects that directly or indirectly have impact on energy issues like: current network traffic [9], speed and stability [8], power consumption peak and valleys [10], power budgeting [24], energy awareness [39]

1.3 *New Design:* which includes studies presenting novel approaches [12], [31], resource management systems [6], simulating software [14], [15], algorithms and mechanisms for new energy-aware framework [17], [18], new methodologies for energy efficiency improvements [13], [16], new kind of usable energy [3], new architecture [34], new developed software [5], [30], [35], [36], [37].

2. *Reports:* This domain includes latest reports on the field of energy and power consumption in cloud computing, focusing on the environmental impact cloud computing has for the moment, its trends for the future, statistics awareness on the area, as well as cloud service providers perspective for the advantages and disadvantages of cloud computing regarding energy consumption.

2.1 *Environmental Impact:* this includes publications that highlight the current state of energy issue on cloud computing, how this current trend can result too problematic regarding environment and worthless consumed energy, as well as new design considerations for green data centers. Main organizations or companies contributing on this sub domain are: Greenpeace [20], [21], [22], U.S. Energy Information Administration [38], Analytic Press [27], Dimension Data [11].

2.2. Cloud Service Provider Perspective: we will mention IBM, Dell and Google Reports. IBM gives its Energy Management Strategy in their Data Centers and IT Efficiency [25]; Dell gives its ideas on usage of techniques such as consolidation and hyperscale hardware, virtualization, integrated solutions for lower power consumption via better hardware designs [40]. Meanwhile Google gives an all new perspective on how cloud computing can in fact be seen as part of the solution towards a more energy efficient future. According to [19], migrating to the cloud can produce estimated energy savings of 68–87% for typical companies and annual energy savings of \$12.3 billion and carbon reductions of 85.7 million metric ton by 2020 for large US companies.

5 Conclusions and discussions

In this paper we gave an overview of the current level of research on cloud computing energy aspects. We presented a literature review classifying a selected group of publications, a total of 40 papers and reports published in the four recent years. Based on a classification scheme, given in section 2, we provided results and analyzed them. We identified that 77.5% of all selected publications are research papers and 22.5% of them are reports. Till 2011' there has been a constant growth in number of publications in both research papers and reports, meanwhile in 2012' we noticed a descent on the number of reports and 2.5 times greater number of paper publications. We justified this as an awareness impact of reports towards researchers of the field.

We also performed a topic-related analysis to show the main current research techniques: hardware or software solutions for better energy-efficiency cloud systems.

To better validate the trend we got from these results, we applied searches of the field in four important international conferences in cloud computing for the years 2010-2012. The results were convincing: years 2011 and 2012 reach a research peak respectively 18 times and 20 times higher than 2010, giving a total of 78 articles on cloud energy area.

Our aim in the future is to be more focused on a specific topic of this research field in order to give more concrete contribution and results.

To conclude, this paper is also an orientation for academics or practitioners interested in the cloud computing area. We encourage them to consider further research on the "energy efficiency in the Cloud" issue as we can see the focus it is given to recently.

References

- 1. Association for Computing Machinery (ACM): Computing Classification System, 2012 Revision
- 2. Alan Roytman, Aman Kansal, Sriram Govindan, Jie Liu, and Suman Nath: Algorithm Design for Performance Aware VM Consolidation, no. MSR-TR-2013-28, 4 March 2013
- Aman Kansal, Bhuvan Urgaonkar, and Sriram Govindan: Using Dark Fiber to Replace Diesel Generators, in XIV Workshop on Hot Topics in Operating Systems (HotOS), USENIX, 13 May 2013
- 4. Aman Kansal, Feng Zhao, Shekhar Srikantaiah: Energy Aware Consolidation For Cloud Computing, Microsoft Research (2012)
- 5. Aman Kansal: Building a More Efficient Data Center from Servers to Software, Microsoft Research, 1 February 2013
- Anton Beloglazov, Rajkumar Buyya: Energy Efficient Resource Management in Virtualized Cloud Data Centers, 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing (2010)

- 7. Anton Beloglazov, Rajkumar Buyya, Young Choon Lee, and Albert Zomaya: A Taxonomy and Survey of Energy-Efficient Data Centers and Cloud Computing Systems, CLOUDS Laboratory, The University of Melbourne, Australia (2011)
- Arka Bhattacharya, Aman Kansal, David Culler, Sriram Sankar, and Sriram Govindan: The Need for Speed and Stability in Data Center Power Capping, in Third International Green Computing Conference (IGCC'12), 5 June 2012
- 9. Christina Delimitrou, Sriram Sankar, Aman Kansal, and Christos Kozyrakis: ECHO-Recreating Network Traffic Maps for Datacenters with Tens of Thousands of Servers, in IEEE International Symposium on Workload Characterization, IEEE, 4 November 2012
- 10.Di Wang, Chuangang Ren, Sriram Govindan, Anand Sivasubramaniam, Bhuvan Urgaonkar, Aman Kansal, and Kushagra Vaid: ACE: Abstracting, Characterizing and Exploiting Peaks and Valleys in Datacenter Power Consumption, in ACM SIGMETRICS, ACM, 17 June 2013
- 11.Dimension Data Report: Seven Design Considerations for a Green Data Centre (2011)
- 12.Drazen Lucanin, Michael Maurer, Toni Mastelic, and Ivona Brandic: Energy Efficient Service Delivery in Clouds in Compliance with the Kyoto Protocol , April 2012
- 13.Dzmitry Kliazovich, Pascal Bouvry, Samee Ullah Khan: DENS: Data Center Energy-Efficient Network-Aware Scheduling, GreenCom Conference (2010)
- 14.Dzmitry Kliazovich, Pascal Bouvry, Samee Ullah Khan: Simulating Communication Processes in Energy- Efficient Cloud Computing Systems, 2012 IEEE 1st International Conference on Cloud Networking (CLOUDNET), Invited Paper
- 15.Dzmitry Kliazovich, Pascal Bouvry, Yury Audzevich, and Samee Ullah Khan: GreenCloud: a packet-level simulator of energy-aware Cloud Computing Data Centers, University of Luxemburg publication (2010)
- 16.Elena Irina Neaga A Digital Infrastructure for Green Utility Computing: The Preliminary Holistic Research Agenda , 3-rd International Conference on Cloud Computing, September 2012
- 17.Eugen Feller, Cyril Rohr, David Margery, Christine Morin: Energy Management in IaaS Clouds: A Holistic Approach, Project-Teams MYRIADS, April 2012
- 18.Eugen Feller, Matthieu Simonin, Anne-Cécile Orgerie and Christine Morin Snooze: An autonomic and energy-efficient management system for virtualized clusters, Myriads, IRISA, Rennes, France, January 2013
- 19.Google Report Google Apps: Energy Efficiency in the Cloud (2012)
- 20.Greenpeace International Report: How Clean Is Your Cloud Catalysing an energy revolution, April 2012
- 21.Greenpeace International Report: How Dirty Is Your Data A Look at the Energy Choices That Power Cloud Computing, April 2011
- 22.Greenpeace International Report: Make IT Green CloudComputing and its Contribution to Climate Change (2010)
- 23.Hadi Goudarzi and Massoud Pedram: Energy-Efficient Virtual Machine Replication and Placement, in a Cloud Computing System (2012)
- 24.Harold Lim, Aman Kansal, and Jie Liu: Power Budgeting for Virtualized Data Centers, in 2011 USENIX Annual Technical Conference (USENIX ATC '11), USENIX, 15 June 2011
- 25.IBM Deutschland Research & Development GmbH Report: Energy Management in a Cloud Computing Environment Smarter Data Centers (2010)
- 26.John R. Savageau: Data Center Consolidation and Cloud Computing: Discussion and Overview, September 2010
- 27.Jonathan G. Koomey: Growth in Data Center Electricity Use 2005 to 2010, Analytics Press Report, August 2011
- 28.Michael Maurer: Towards Energy-efficient Cloud Computing , VITALAB Distributed Systems Group (2012)

- 29.Monika Malhothra, Shuchi Sharma: A True Green Cloud Data Centers, International Journal of Computing & Business Research, 'I-Society 2012' Proceedings (2012)
- 30.Nitin Singh Chauhan, Ashutosh Saxena: A Green Software Development Life Cycle for Cloud Computing, IEEE Computer Society, February 2013
- 31.Peter Xiang Gao, Andrew R. Curtis, Bernard Wong, S. Keshav: It's Not Easy Being Green, SIGCOMM'12
- 32.Pooja. R.Kalange, Prashant Vyankatrao Raut: Green Cloud Computing: Energy Efficiency and Environmental Sustainability, International Journal of Science, Engineering and Technology Research (IJSETR), December 2012
- 33.Rajkumar Buyya: Energy-Efficient Cloud Computing: Opportunities and Challenges, ManjraSoft (2012)
- 34.Rayman Preet, S. Keshav, Tim Brecht: HITS: An Cloud-based architecture for energy management in the Smart Grid , Information Systems and Science for Energy (ISS4E) Laboratory, March 2012
- 35.Ripal Nathuji: GreenClouds: Combining cloud computing software with energy efficient ARM hardware , CloudStack Collaboration Conference (2012)
- 36.Sriram Govindan, Jie Liu, Aman Kansal, and Anand Sivasubramaniam: Cuanta: Quantifying Effects of Shared On-chip Resource Interference for Consolidated Virtual Machines, in ACM Symposium on Cloud Computing (SOCC), ACM, 27 October 2011
- 37.Sriram Govindan, Jie Liu, Aman Kansal, and Anand Sivasubramaniam, Cuanta: Quantifying Effects of Shared On-chip Resource Interference for Consolidated Virtual Machines, no. MSR-TR-2011-55, May 2011
- 38.U.S. Energy Information Administration (EIA) Report: International Energy Outlook 2011, September 2011
- 39.Weiwei Xiong and Aman Kansal: Energy Efficient Data Intensive Distributed Computing, in IEEE Data Engineering (Special Issue on Energy Aware Big Data Processing), IEEE Computer Society, 31 March 2011
- 40.Yoram Heller: Finding Energy Efficiency in the Cloud: Why it Matters, and How to Get There, January 2013

Improved MDS-MAP Algorithm for Nodes Localization in 3D Wireless Sensor Networks

Biljana Risteska Stojkoska Faculty of Computer Science and Engineering University "Ss. Cyril and Methodius" Skopje, Macedonia biljana.stojkoska@finki.ukim.mk

Abstract. In the recent years, there has been huge advancement in wireless sensor computing technology. Today, Wireless Sensor Network (WSN) has become a key technology for different types of smart environment. Nodes localization in WSN has arisen as a very challenging problem in the research community. Most of the applications for WSN are not useful without a priory known nodes positions. Adding GPS receivers to each node is an expensive solution and inapplicable for indoor environments. In this paper, we implemented and evaluated improved MDS-MAP algorithm for three dimensional (3D) nodes localization in WSN. Using extensive simulations we investigated our approach regarding various network parameters. We compared the results from the simulations with other approaches for 3D-WSN localization and showed that our algorithm outperforms other techniques in terms of accuracy.

Keywords: Wireless Sensor Network, Multidimensional Scaling, 3D Localization

1 Introduction

Wireless Sensor Network (WSN) represents a collection of wireless sensor nodes that coordinate with each other in order to perform a particular task. Sensor node is low cost and low power device that consists of three components: a sensing subsystem for data acquisition, a processing subsystem for local data processing and a wireless communication subsystem for data transmission[1][2]. Data are transferred from sensor nodes to the sink node (base station) through a multi-hop communication paradigm. Each sensor sends data to its closest neighbor responsible for retransmitting the packets [3].

Following the latest developments in computer and communication technologies, everyday objects are becoming smarter, as ubiquitous connectivity and modern

sensors allow them to communicate with each other. The deployment of sensors and actuators everywhere around us adds a new dimension to the world of information and communication, which enables the creation of new and enriched services widely applied in different industrial and civilian application areas, including industrial process monitoring and control, machine health monitoring, environment and habitat monitoring, healthcare applications and traffic control [4][5].

A fundamental problem in wireless sensor networks is localization i.e. the determination of the geographical locations of sensors. The most straightforward solution to the localization problem is to apply global positioning system (GPS) to each node. But it is not an attractive solution because of cost, size and power constraints [6][7][8]. Thus, an effective localization algorithm should employ all the available information from the nodes to compute all the positions.

Localization is a challenge when dealing with wireless sensor nodes, and a problem which has been studied for many years. Many different techniques have been proposed for solving this problem, but most of them consider only two-dimensional (2D) network. Hence, localization issue in three dimensions remains a challenging problem in the research community.

In this paper, we investigate classical multidimensional scaling technique (MDS-MAP) for nodes localization in three dimensional WSN and propose improved MDS (IMDS) algorithm that use a heuristic approach for distance matrix calculation. IMDS improves the accuracy compared with well known MDS-MAP. The rest of this paper is organized as follows. In the second section, the relevant work related to the present 3D localization techniques is discussed. The third section refers to multidimensional scaling as a technique for nodes localization WSN. The fourth section elaborates IMDS. Section five presents the results provided from our simulations. Finally, we conclude this paper in section six.

2 Related work

Techniques for WSN localization can be basically divided into two categories: rangebased and range-free methods. The first are considered more accurate and most of the algorithms for localization belong to this category [6][9]. They use the distance between the nodes in the network [10][11]. RSSI (Receive Signal Strength Indicator) is the most common technique used for distance estimation. RSSI utilizes small resources without the need for extra hardware. RSSI measures the power of the received radio signal to calculate the distance between two nodes that are in each other transmission range. Other techniques (Time of Arrival - ToA, Time Difference of Arrival – TDoA, etc.) for distance measurement translate propagation time into distance. This can be done if signal propagation speed is known in advance. These techniques can be used with acoustic, infrared and ultrasound signals.

Many research groups have investigated different techniques for nodes localization in WSN, but most of the proposed techniques consider only two-dimensional WSN

localization. A few researchers within the last years have tried to focus on threedimensional localization.

One of the first proposed techniques for 3d localization is Landscape-3d [12]. In the first phase, location-unaware nodes measure a set of distances to mobile location assistants (LAs) using AOAs and RSSI. In the next phase, nodes use Unscented Kalman Filter to estimate their own position. Very similar approach is proposed in [13]. RSSI is used for distance measurements while particle filter for node positioning.

Although these methods are independent of node densities and network topologies, the major drawback is its dependence on mobile beacon that might not be available under some deployments (e.g. hostile environments).

In [14] the authors propose cluster-based approach named CBLALS. In each cluster, the inter-cluster range measurement errors are corrected using triangle principle. The evaluation of CBLALS with respect to DV-Distance (3D) shows that CBLALS has much better positioning accuracy than DV-Distance (3D).

A Novel Centroid localization method that significantly improves the basic centroid localization algorithm is presented in [15]. Each unknown node randomly selects four anchor nodes (nodes whose coordinates are known in advance) in range to form a series of tetrahedrons used to calculate its own position with this Novel Centroid method.

3 Multidimensional scaling for localization in 3D-WSN

Multidimensional scaling (MDS) is a set of techniques used for reducing the dimensionality of the data (objects). MDS visualize the results in order to show hidden structures in the data [16]. MDS algorithm uses the distances between each pair of object as input and generates 2D-points or 3D-points as output.

In WSN, MDS has an ability to reconstruct the relative map of the network even without anchor nodes [17][18]. The first and most explored is MDS-MAP [18] that is very accurate for density networks. More computationally dependent approaches are introduced in [19][20]. MDS-MAP(P) [19] is a modification of the MDS-MAP that computes local maps at each node in the network and then integrates them into a global map. One different approach based on MDS-MAP is introduced in [21] and [22]. It is known as cluster-based MDS, where the network is divided into clusters responsible for partial localization using MDS-MAP. This approaches show better performances than MDS-MAP and MDS-MAP(P) for irregular network topologies in terms of accuracy and computational complexity.

Consider a WSN with n nodes in a three dimensional space. Let S_{nx3} be an unknown matrix where each row represents the coordinates of i-th point (node) along three dimensions. Let $\Delta^{(2)}(S)$ or $\Delta^{(2)}$ represents the matrix of squared Euclidean

distances between nodes i and j (i=1..n, j=1..n). The $\Delta^{(2)}$ matrix can be calculated since we assume that sensors (nodes) have mechanisms to estimate the distance between each other.

$$\Delta^{(2)}(S) = \Delta^{(2)} = \begin{bmatrix} 0 & \Delta^{(2)}_{12} & \Delta^{(2)}_{13} & \dots & \Delta^{(2)}_{1n} \\ \Delta^{(2)}_{21} & 0 & \Delta^{(2)}_{23} & \dots & \Delta^{(2)}_{2n} \\ \dots & & & \\ \Delta^{(2)}_{n1} & \Delta^{(2)}_{n2} & \Delta^{(2)}_{n3} & \dots & 0 \end{bmatrix}$$
(1)

where

$$\Delta^{(2)}_{ij}(S) = \Delta^{(2)}_{ij} = \sum_{a=1}^{3} (s_{ia} - s_{ja})^2$$
⁽²⁾

It can be rewritten as:

$$\Delta^{(2)}{}_{ij} = \sum_{a=1}^{3} (s_{ia}{}^{2} + s_{ja}{}^{2} - 2s_{ia}s_{ja})$$
(3)

or
$$\Delta^{(2)} = c1' + 1c' - 2SS'$$
 (4)

where, 1 is an n x 1 vector of ones, c is a vector consisting the diagonal elements of the scalar product matrix, i.e., $c = \sum_{a=1}^{3} s_{ia}^{2}$.

Multiplying both sides of (4) by centering matrix T

$$T = I - n^{-1} 11', (5)$$

where I is the identity matrix, we get:

$$T\Delta^{(2)}T = T(c1' + 1c' - 2SS')T$$

$$T\Delta^{(2)}T = Tc1'T + T1c'T - T2SS'T$$

$$T\Delta^{(2)}T = Tc1'T + T1c'T - T(2B)T$$

$$T\Delta^{(2)}T = -T(2B)T$$

$$B = -\frac{1}{2}T\Delta^{(2)}T$$
(6)

Since B is symmetric it can be decomposed:

$$B = Q\Lambda Q' = (Q\Lambda^{1/2})(Q\Lambda^{1/2})' = SS'$$
(7)

where Q is orthonormal and Λ is a diagonal matrix.

$$\Rightarrow S = Q\Lambda^{1/2} \tag{8}$$

The recovered matrix S is rotated and has a different coordinate system than the original S.

MDS-MAP for 3D WSN consists of 3 steps:

1. Calculate shortest distances between every pair of nodes (using either Dijkstra's or Floyd's all pairs shortest path algorithm). This is the distance matrix that serves as input to the multidimensional scaling in step 2.

2. Apply classical multidimensional scaling to the distance matrix. The first 3 largest eigenvalues and eigenvectors give a relative map with relative location for each node.

3. Transform the relative map into absolute map using sufficient number of anchor nodes (at least 4). This process usually includes translation, rotation and reflection.

3.1 Finding optimal rotation and translation between corresponding 3D nodes

Generating an absolute map (step 3) of the WSN requires anchor nodes. At least four sensors' physical positions are needed in order to identify the physical positions of remaining nodes in the group in 3D case.

Let $P = \{p_1, p_{2,...,} p_N\}$ and $Q = \{q_1, q_{2,...,} q_N\}$ be two sets of corresponding nodes, where N is the number of anchor nodes in the WSN. We wish to find a transformation that optimally aligns the two sets in terms of least square errors, i.e., to minimize the sum of squares of the errors between estimated positions of the anchors from MDS map and their true positions. We seek a rotation matrix **R** and a translation vector **t** such that

$$(R,t) = \arg\min_{R,t} \sum_{i=1}^{N} \left\| (Rp_i + t) - q_i \right\|^2,$$
(9)

This transformation is also known as Euclidean or Rigid transformation, because it preserves the shape and the size.

There are many algorithms purposed in the literature that compute a rigid 3D transformation. The most explored are based on:

- Singular Value Decomposition (SVD),
- Unit Quaternion (UQ)
- Dual Quaternion (DQ)
- Orthonormal Matrices (OM)

A comparison of these four methods can be found in [23]. It is shown in [24] that the results of all these methods are similar in most cases and the difference in accuracy is almost insignificant, but the SVD is the most stable.

Finding the optimal rigid transformation with SVD can be broken down into the following steps:

1. Compute the weighted centroids of both point sets

$$\overline{p} = \frac{1}{N} \sum_{i=1}^{N} p_i, \ \overline{q} = \frac{1}{N} \sum_{i=1}^{N} q_i$$

- 2. Compute the centered vectors $p_i := p_i \overline{p}, q_i := q_i \overline{q}, i=1,...,N$.
- 3. Compute the 3x3 covariance matrix $H = P'Q'^T$, where P' and Q' are the 3xN matrices that have p_i and q_i as their columns, respectively.
- 4. Compute the singular value decomposition $H = U\Sigma V^T$. The rotation we are looking for is then $R = VU^T$.
- 5. Compute the optimal translation as $t = \overline{q} R\overline{p}$.

3.2 Time complexity of MDS-MAP for 3D-WSN

In step 1, distance matrix construction using Dijkstra's or Floyd's algorithm requires $O(n^3)$, where n is the number of nodes in the network. In step 2, applying MDS to the distance matrix has complexity of $O(n^3)$ due to singular value decomposition. In step 3, the relative map is transformed through linear transformations. Computing the rigid transformation takes O(N) time for computing P and Q, while computing SVD takes only $O(3^3)$ time (since the covariance matrix H has dimension 3x3). Applying the transformation (rotation and translation) to the whole relative map takes O(n-N) time, where N is the number of anchors (N<<n).

4 Improved MDS (IMDS)

The main drawback of MDS-MAP is the way it calculates the distance matrix. Using Dijkstra (or Floyd's) all pairs shortest path algorithm to estimate the distances between non-neighboring nodes in the network gives incorrect distance matrix. Dijkstra distance between two nodes usually correlates with the Euclidean distance, but always calculates the longest possible distance.

To reduce the error present in distance matrix, we introduce a heuristic approach to estimate the distances between non-neighboring nodes.

4.1 Heuristic approach for distance matrix calculation (HA)

Consider there are three nodes in a network: A, B and C (Fig. 1), with known distances between nodes A and B (d_1 =AB), and between nodes B and C (d_2 =BC). Since distance matrix requires the distance between each pair of nodes in the network, the distance between nodes A and C have to be calculated (Fig. 1).

If maximum radio range of the nodes in the network is R, than we know for sure that node C lies on the curve $C_1 C_2$. If Dijkstra's algorithm is used for this purpose, it will approximate AC=AB+BC, which is the longest possible theoretical distance between nodes A and C. On the other hand, if we calculate the shortest possible theoretical distance between distance between nodes A and C, it will corresponds to C_1 (AC=R).

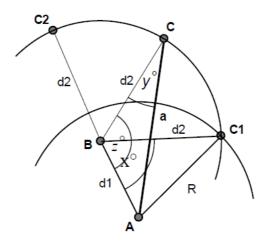


Fig. 1. Distance approximation

To minimize the possible error, we purpose heuristic solution that assumes that the node C lies exactly in the middle of the curve $C_1 C_2$. Hence, the distance *a*=AC can be calculated using cosine formula as:

$$a^{2} = d_{1}^{2} + d_{2}^{2} - 2 \cdot d_{1} \cdot d_{2} \cdot \cos{(\blacktriangleleft ABC)}$$

where d_1 and d_2 are known and $\blacktriangleleft ABC$ can be easily calculated using cosine formula.

IMDS for 3D WSN consists of the following 3 steps:

1. Calculate shortest distances between every pair of nodes using heuristic approach. This is the distance matrix that serves as input to the multidimensional scaling in step 2.

2. Apply classical multidimensional scaling to the distance matrix. The first 3 largest eigenvalues and eigenvectors give a relative map with relative location for each node.

3. Finding the optimal rigid transformation with SVD and transform the relative map into absolute map using sufficient number of anchor nodes (at least 4).

IMDS preserves the time complexity of MDS-MAP algorithm.

5 Results and discussion

We assume a typical sensor network composed of hundreds (or thousands) of sensor nodes deployed uniformly across a three dimensional monitored area. Each sensor is equipped with an omni-directional antenna and only nodes within certain radio range R can communicate with each other. We made following assumptions:

- Nodes are static and unaware of their location.
- There is a path between every pair of nodes.
- Nodes deployed in close proximity to each other exchange messages.
- Each node uses RSSI (or any other) method for distance estimation.

We simulated both IMDS and MDS-MAP on random network topology with Matlab. 100 nodes were uniformly placed in a cubic area (100r x 100r x 100r, where r is a unit length distance). Our work was mainly focused on network properties like number of anchors, average connectivity and range error. We consider:

- Different network topologies:
 - o random deployment for 3D-WSN (100 nodes)
 - o grid deployment for 3D-WSN (125 nodes).
- Different number of anchors (4, 6, 10 and 15) for absolute map construction. In our experiments we use SVD method for 3D rigid transformation. In our simulation, the anchors were selected randomly.
- Different radio ranges (R) which lead to different average connectivity (average number of neighbors).
- Different radio range error *er* (from 0 to 30% of R with step 5% of R)

Thus 280 different networks were simulated $(2 \times 4 \times 5 \times 7)$ and each node location was discovered with both MDS-MAP and IMDS technique. The connectivity parameter and the estimation error for each scenario represent average over 30 trials. The average estimation error is normalized by the radio range R:

$$Error = \frac{\sum_{i=1}^{n} \text{distance}(pos_{i}^{(estimated)} - pos_{i}^{(true)})}{(n-N) \cdot R} \cdot 100\%$$
(10)

where n is the number of nodes in the network, N is the number of anchor nodes, $pos_i^{(estimated)}$ is the estimated location and $pos_i^{(true)}$ is the true location of the i-th node.

5.1 Comparison of HA and Dijkstra

In order to evaluate our heuristic approach (HA) for distance matrix calculation, we calculated the distances between the nodes using both HA and Dijkstra algorithm. Then we compared the differences between the real distances and the distances obtained using the above-mentioned algorithms. The difference is normalized with radio range R. The results from the simulation are presented on Fig. 2.

As can be seen from the figure, HA performs better than Dijkstra for all connectivity levels. This is expected knowing that \underline{D} ijkstra calculate the longest possible distance, while HA tend to shorten this distance.

HA performs much better than Dijksta especially for large range error *er*. This is rather important characteristic of HA since range measurement in the real applications is prone to error. When RSSI is used for distance signalization, the range error measurement is at least 10% R. The results presented in [25] show average range error measurement between 5% R and 30% R for longer radio range R. Similar research conducted in [26] and [27] that investigate RSSI reported average measurement error around 20% R.

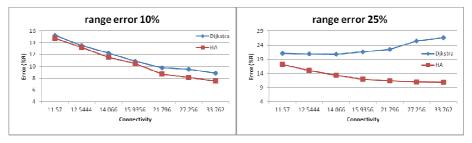


Fig. 2. Distance error for HA and Dijkstra

5.2 Comparison of MDS-MAP and IMDS for 3D WSN

Fig. 3 shows an example of typical 3D network with 100 nodes randomly deployed (R=35r and an average connectivity of 11.6). Blue lines represent the distance between the absolute and the estimated position when using IMDS and MDS-MAP algorithm respectively. The estimation error is larger if the lines are longer. The absolute map is achieved using 10 anchors (red circles). As can be seen from the figure, IMDS performs better than MDS-MAP.

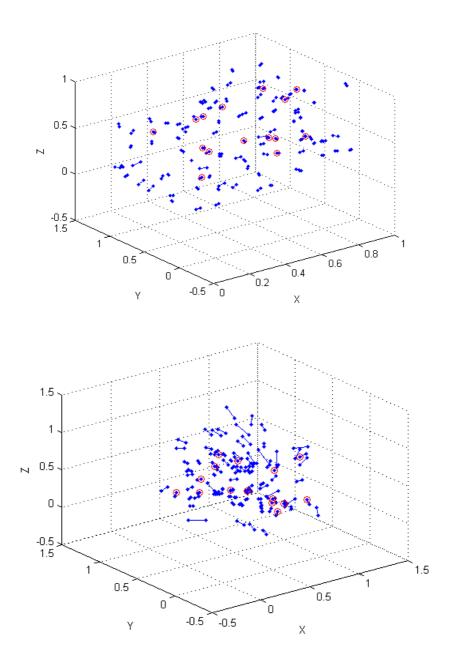


Fig. 3. Estimation error for IMDS (upper) and MDS-MAP (lower)

Fig. 4 shows the results of MDS-MAP and IMDS for random topology with range error 10% of R and 25% of R. As can be seen from the figure, in both cases IMDS performs smaller estimation error than MDS-MAP for all connectivity levels.

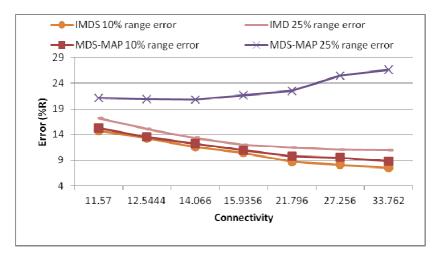


Fig. 4. Estimation error for IMDS (upper) and MDS-MAP (lower)

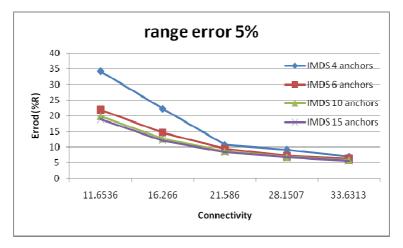


Fig. 5. The effect of number of anchors on the estimation error

We must note that range error *er* have great impact on estimation accuracy. When range error *er* is small, MDS-MAP and IMDS perform almost the same, with IMDS being slightly better. As *er* increases, MDS-MAP rapidly deteriorates, while IMDS is pretty much stable. Considering the fact that range error is inevitable phenomena in WSN, we can conclude that IMDS is better option when choosing localization algorithm for 3D environments.

As expected, using more anchor nodes reduces the estimation error, but has no significant impact. Number of anchors affects the results when the connectivity level is low. For high connectivity levels, there is no evident improvement (Fig. 5).

5.3 Comparison of MDS-MAP and IMDS with other approaches for 3D localization

To evaluate IMDS more convincingly, the simulation experiments were conducted for CBLALS method [14] and for the Novel Centroid Algorithm [15]. The comparison with [14] shows that both IMDS and MDS-MAP perform better than CBLALS in terms of localization accuracy (Fig. 6 upper). It is assumed that the density of the anchors is 10%, and the connectivity of the networks is 10~15. As can be seen from the results, MDS-MAP is much more robust to range error than CBLALS, especially for large range errors. The result from the comparison of IMDS and MAP-MAP with the Novel Centroid Algorithm is shown on Fig. 6 (lower).

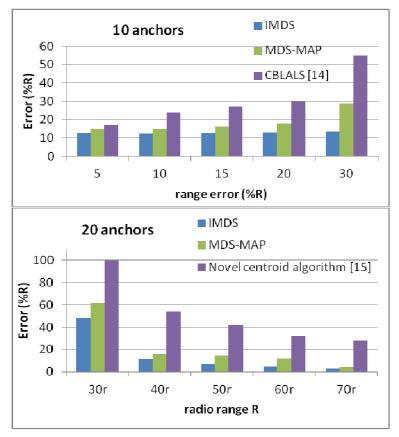


Fig. 6. Comparison of IMDS and MDS-MAP with CBLALS[14] and Novel Centroid algorithm[15]

For all different radio range R, both IMDS and MDS-MAP performs better than [15]. Here the density of the anchors is 20%. The radio range error for this simulation is er=0%R.

6 Conclusion

In this paper, we presented improved MDS-MAP algorithm for nodes localization in 3D-WSN. Through extensive simulations we evaluated the algorithm and showed that IMDS outperforms other approaches presented in [14] and [15] in terms of accuracy.

For future work, we intend to investigate our algorithm on irregular three-dimensional network topologies, where nodes are deployed on more complex 3D terrains, like terrain with a valley or terrain with a mountain. It is expected that MDS-based algorithms for WSN localization will not work well for such scenarios, basically because of multi-hop distance between each pair of nodes. We conducted initial simulations which prove this assumption. For the future work, we will extend our algorithm considering hierarchical network organization based on cluster formation. This cluster-based approach which is already developed and implemented for 2D networks in [22] encourages us to consider cluster-based extension for 3D networks.

References

- 1. Akyildiz, I.F., Weilian, S., Sankarasubramaniam, Y., Cayirci, E.: A Survey on Sensor Network. IEEE Communications Magazine 40, 102–114 (2002)
- F.L.Lewis, Wireless Sensor Networks, chapter 4 in D. J. Cook and S. K. Das, editors, Smart Environments: Technologies, Protocols, and Applications, John Wiley, New York, 2004;
- 3. K. Akkaya, M. Younis, A survey of routing protocols in wireless sensor networks, Elsevier Ad Hoc Network Journal, 3(3):325–349, May 2005.
- 4. Puccinelli D, Haenggi H. 2005. Wireless sensor networks: applications and challenges of ubiquitous sensing. *IEEE Circuits and Systems Magazine* 2005 (3): 19–29.
- Garcia-Hernandez, C. F., Ibarguengoytia-Gonzalez, P. H. and Perez-Diaz, J. A. (2007) 'Wireless Sensor Networks and Applications: A Survey', *IJCSNS International Journal of Computer Science and Network Security*, 7(3), 264-273.
- 6. J. Wang, R.K. Ghosh, and S.K. Das. "A survey on sensor localization". Journal of Control Theory and Applications, 8(1):2-11, 2010.
- 7. Pal. Localization Algorithms in Wireless Sensor Networks: Current Approaches and Future Challenges. In *Network Protocols and Algorithms*, 2010, Vol. 2, No. 1
- 8. Rudafshani, M., Datta, S. Localization in wireless sensor networks, In: 6th international conference on Information processing in sensor networks, pp. 51-60 (2007)
- T. He, C. Huang, B. Blum, J. Stankovic, and T. Abdelzaher, "Range-free localization schemes in large scale sensor networks", In *Proceedings of the Ninth Annual International Conference on Mobile Computing and Networking (MobiCom'03)*, September 2003, San Diego, CA, USA, pp. 81-95.
- G. Mao, B. Fidan, B. D. Anderson. Wireless sensor network localization techniques [J]. Computer Networks, 2007, 51(10): 2529 – 2553.

- 11. R. I. Reza, "Data fusion for improved TOA/TDOA position determination in wireless systems," Ph.D. Dissertation, Virginia Tech., 2000.
- 12. Xiaobo Zhou, Liqiang Zhang and Qiang Cheng. (2006) "Landscape-3d: A robust localization scheme for sensor networks over complex 3d terrains." In *Proc. 31st IEEE International Conference Local Computer Networks*, pages 239--246, November, 2006.
- 13. Caballero, F.; Merino, L.; Maza, I.; Ollero, A. A Particle Filtering method for Wireless Sensor Network Localization with an Aerial Robot Beacon. In *IEEE International Conference on Robotics and Automation*, Pasadena, CA, USA, 2008; pp.596–601.
- J. Shu, R. Zhang, L. Liu, X. Wu, Z. Zhou, "Cluster-based Three-dimensional Localization Algorithm for Large Scale Wireless Sensor Networks," Journal of Computers, vol 4, no 7 (2009), 585-592, Jul 2009.
- H. Chen, P. Huang, M. Martins, H. C. So, K. Sezaki, Novel centroid localization algorithm for three-dimensional wireless sensor networks, in Proc. 4th International Conference on WirelessCommunications, Networking and Mobile Computing (WiCOM 2008), Oct. 2008, Dalian, China.
- 16. T. Cox and M. Cox, Multidimensional Scaling, Chapman & and Hall, London, 1994.
- 17. Xiang Ji and Hongyuan Zha, "Sensor Positioning inWireless Ad-hoc Sensor Networks using Multidimensional Scaling," *Proceedings of 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2004)*, March 2004.
- 18. Yi Shang, Wheeler Ruml, Ying Zhang, and Markus P. J. Fromherz, "Localization from mere connectivity," in *Proceedings of ACM International symposium on mobile ad hoc networking and computing*, pp. 201-212, June 2003.
- 19. Yi Shang and Wheeler Ruml, "Improved mds-based localization," in *Twenty-third* AnnualJoint Conference of the IEEE Computer and Communications Societies (INFOCOM), pp. 2640–2651, 2004
- V. Vivekanandan and V. W. Wong. Ordinal mds-based localization for wireless sensor networks. In VTC-2006: IEEE 64th Vehicular Technology Conference, pages 1-5, 2006.
- Gwo-Jong Yu, and Shao-Chun Wang "Hierarchical MDS-based Localization Algorithm for Wireless Sensor Networks", 22nd International Conference on Advanced Information Networking and Applications, 25-28 March, 2008, pp.748-754.
- 22. Min Han Shon, Minho Jo, Hyunseung Choo: An interactive cluster-based MDS localization scheme for multimedia information in wireless sensor networks. Computer Communications 35(15): 1921-1929 (2012)
- 23. Eggert, D. W., Lorusso, A., and Fisher, R. B. "Estimating 3-D Rigid Body Transformations: A Comparison of Four Major Algorithms," Machine Vision and Applications (1997) Vol. 9, No. 5/6, pp. 272–290.
- Lorusso A., Eggert D., Fisher R.: "A Comparison of Four Algorithms for Estimating 3-D Rigid Transformations", In: Proceedings of the 4th British Machine Vision Conference (BMVC 1995), Birmingham, England, pp. 237-246, (September 1995).
- 25. Z. Jianwu, Z. Lu, "Research on distance measurement based on RSSI of ZigBee," Computing, Communication, Control, and Management, 2009. CCCM 2009. ISECS International Colloquium on, vol.3, no., pp.210-212, 8-9 Aug. 2009.
- A. Awad, T. Frunzke, F. Dressler, "Adaptive distance estimation and localization in WSN using RSSI measures," in Proc. 10th Euromicro Conf. Digital Syst. Des. Architectures, Methods Tools (DSD) 2007, pp. 471–478.
- A. Faheem, R. Virrankoski, M. Elmusrati, "Improving RSSI Based Distance Estimation for 802.15.4 Wireless Sensor Networks," in Conference Proceedings of the 2010 IEEE International Conference on Wireless Technology and Systems (ICWIT 2010), Aug. 2010.

Wireless Sensor Networks Framework for Indoor Temperature Regulation

Biljana Stojkoska¹ and Andrijana Popovska Avramova²

¹ Faculty of Computer Science and Engineering, University Ss. Cyril and Methodius, Skopje, Macedonia biljana.stojkoska@finki.ukim.mk ² Department of Photonics Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark apop@fotonik.dtu.dk

Abstract. Wireless Sensor Networks take a major part in our everyday lives by enhancing systems for home automation, health-care, temperature control, energy consumption monitoring etc. In this paper we focus on a system used for temperature regulation for homes, educational, industrial, commercial premises etc. We propose a framework for indoor regulation and optimization of temperature using wireless sensor networks based on ZigBee. Methods for optimal temperature regulation are suggested and discussed. The framework is based on methods that provide energy savings by reducing the amount of data transmissions through prediction methods. Additionally the framework explores techniques for localization, such that the location of the nodes is used for optimization of the temperature settings. Information on node location is used to provide the most optimal tradeoff between the time it takes to reach the desired temperature at a specific part of the room and energy consumption.

Keywords: temperature optimization, wireless sensor networks, ZigBee

1 Introduction

Wireless sensor networks (WSN) are able to efficiently sense various information with high accuracy and low power consumption. The development of sensors and networks based on sensors has impacted and changed everyday life and work. Engaging WSN in home and industrial monitoring systems, medicine and health-care systems, entertainment, education etc. has enlighten and improved processes in overall. A wireless sensors consist of three major elements [1]: sensor unit (used to measure the parameter), computing unit (used to process data) and communication unit (based on radio communication). Different radio technologies can be used such as: ZigBee, Wi-Fi, Bluetooth, GSM etc. ZigBee as an emerging technology has been proven to make WSN self-configurable and self-healing while operating at low power consumption [2].

Intelligent smart home frameworks have been proposed recently by the research community in [3] and [4]. The proposed systems are used to monitor and report different home parameters such as: temperature, humidity, light and in order to control different electrical devices: lightning, aircondition, heaters. In [4] energy optimization is based on a dynamic programming algorithm that will control the usage of energy and sell it back to the smart grid.

This paper proposes a Wireless Sensor Network Framework for Indoor Temperature Regulation (WSN-FITR). Homes, classrooms and halls are often heated up by a number of temperature controlled heaters. Users are usually not interested in controlling the temperature at each separate heater. The radiators for examples are normally located just above the floor/below windows and at the room's walls and furthermore the measurements don't show the real room temperature as the temperature sensors are located just next to the heater. Neighbouring rooms with own heating elements also influence the temperature in the controlled room. This paper shows how the node localization methods can be used for room temperature optimization in order to provide the most optimal tradeoff between the time it takes to reach the wanted temperature at a specific part of the room and energy consumption. The distance from the heaters to other networks and their temperature influence is important for saving the overall energy consumption. Even more, reduction of the required data transmission through prediction methods is considered. This is important in order to increase the battery life of the nodes and to extend the network lifetime.

The rest of this paper is organized as follows. In the second section, the framework architecture of WSN-FITR is presented. Section 3 gives a detailed explanation of techniques for indoor localization and clustering with respect to characteristics of the environment where WSN is deployed. Section 4 compares techniques for data reduction in WSN. Finally, we conclude this paper in the last section.

2 System Architecture

The proposed WSN-FITR system for indoor temperature regulation is consisted of two basic elements: sensor-regulators and temperature controllers which are inter-connected in a ZigBee network. In wireless sensor network the sensorsregulators are known as nodes and the controller is referred to as sink nodes. Deployment diagram of the framework is given on Fig. 1.

Sensor-regulators perform measurement and reporting local temperature readings to the controller. They are attached to a device and regulate its action (for example increase/decrease heating) in order to reach a certain temperature. Typical devices and their corresponding actions are:

- **Heating bodies** such as central heating radiators, electric radiators or fans. The regulator can increase/decrease the heating by using the valve controller.
- Air conditioners which can be based on a fan. The regulator can increase/decreasing the cooling volume in order to reach a certain temperature.
- Air flow such as central air flow ventilation. These devices can regulate the degree of air circulation.

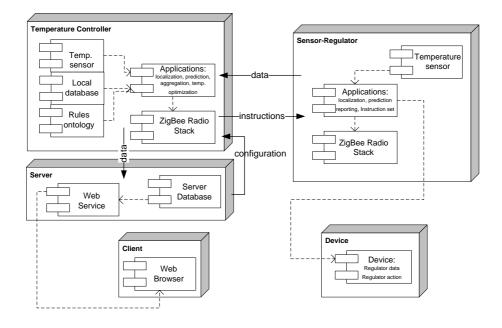


Fig. 1. Deployment diagram of the system.

• Window shutters such as outside curtains. By rising/lowering the curtains the influence of the sun energy can be regulated.

The temperature controller nodes are considered more powerful than the sensor-regulators. They are expected to have more advanced capabilities: memory, processing unit and steady energy supply. They should be located higher up in the room and away from all the heaters and windows in a location that better represents the room's temperature in order to measure the most relevant temperature. Additionally they are used to control the temperature at the premises where the device is placed by generating instructions for regulation. Temperature controllers represent the sink element where the information is gathered and locally analysed. Two different types of information are considered static and dynamic. Static data is related to one-time information such as location (in 3D) and type of nodes. Dynamic data is related to time variable parameters such as temperature, energy cost etc.

As illustrated on Fig. 1, the temperature controller consists of the following elements: local database, rules (which can be based on ontologies [5]). Client should be able to remotely configure the rules so the controller can meet the temperature goals. The rules can include for example temperature levels for several time intervals during different days. The local database at the controller can be used in order to store information regarding the temperature readings, as well as the static information such as node type and locations. The database can further contain information on the size of the room being monitored, the number of expected visitors and other factors that can influence the temperature changes. All information that can not be obtained from the nodes, such as electricity cost, can be obtained from the server. Additionally the server contains database that can be used by the controller to store historical data.

2.1 Temperature Optimization Framework

Finally a temperature controller needs to make decisions by generating a set of instructions for the different nodes by running a regulation method. This method is required by the controllers in order to achieve distributed decision regarding the temperature difference that need to be achieved in a certain point in order to satisfy the required overall temperature.

In order to provide a self-organized and cost-effective solution, the WSN must provide the following functionalities: self-localization, nodes clustering, data prediction, distributed decision making etc. Sensor readings are useless if the location where they are measured is not known. Thus, a suitable localization algorithm should be implemented in order to discover the location of the nodes. This is important because manual recording of the nodes positions is very time consuming solution and prone to errors. The location information will be used by the controller device to deduct the temperature set-point that each of the heater shall be commanded to, so that the temperature dissipation from all heater gives the wanted temperature at the controller's location and the controlled premises in overall (elaborated in section 3). After nodes discovering phase, nodes can be divided into clusters. There are many algorithms purposed in the literature for optimal nodes clustering. For now we assume that nodes deployed in close proximity to each other belong to the same cluster.

After these two phases (discovering and forming the clusters), we consider that WSN is established. Nodes can start measuring the temperature and forwarding the measured readings to the final destination (sink node). In order to save energy, algorithms for data prediction should be implemented on both sides: node and sink (elaborated in section 4).

The temperature measurements are analysed by generation of temperature gradients. The temperature gradient indicates the direction and the rate at which the temperature changes within a particular location. The dynamic information such as temperature gradients, time of the day, expected visitors, electricity cost and the static information such as node type represent input to the temperature optimization method.

The location and temperature prediction methods allow the calculation of the temperature gradient. Several methods can be used in order to rank or assign weights to the different types of input such as Multiple Attribute Decision Making algorithm, Genetic Algorithms, Analytic Hierarchy Process, Fuzzy logic. Fuzzy logic is suitable as fuzzy judgement matrices used for the comparative analysis are close to the way the humans reflect and are very easy to implement. The Fuzzy Inference System can be based on the standard Mamdani or Sugeno. For each parameter that needs to be taken into consideration membership function needs to be defined. Additionally the rules that need to be applied need to be carefully chosen. The actual definition of the membership functions and the rules will be considered in our future work.

After the collected data is analysed, the controller sends instructions to the nodes (regulators). The instructions are represented by the temperature difference that the node should achieve. If the temperature needs to be increased, the nodes perform an action such as increasing the heating energy.

2.2 WSN-FIRT network topology

ZigBee is a low-cost, low-power, wireless mesh network standard. The low cost allows the technology to be widely deployed in wireless control and monitoring applications (in fields such as home automation, health-care, temperature control etc.). Low power-usage allows longer life with smaller batteries. Mesh networking provides high reliability and more extensive range. ZigBee operates in the industrial, scientific and medical (ISM) radio bands; 868 MHz in Europe, 915 MHz in the USA and Australia and 2.4 GHz in most jurisdictions worldwide. Data transmission rates vary from 20kbps in the 868 MHz frequency band to 250kbps in the 2.4 GHz frequency band.



Fig. 2. Star architecture for a typical classroom.

The ZigBee network layer natively supports both star and tree network topologies, and generic mesh networks. Every network must have one coordinator device, tasked with its creation, the control of its parameters and basic maintenance. Within star networks, the coordinator must be the central node. Both trees and meshes allows the use of ZigBee routers to extend communication at the network level. Since the routers need to be constantly on listening for network traffic, it is normally assumed that the routers are mains powered sensors/devices and the battery devices are assumed to be sleeping and only waking up and polling for data periodically or on demand (upon user interaction).

The network topology should adapt to the characteristics of the controlled premiss. For the proposed WSN-FITR we consider two different topologies: star and cluster-based. An example of a star topology is illustrated in Fig. 2 where one classroom is illustrated. In this case there is a single temperature controller that is responsible for controlling the temperature at the different nodes. The

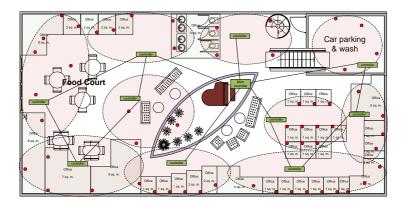


Fig. 3. Clustered architecture for one floor in a shopping center.

star topology is most appropriate for small areas where there are no major obstacles so that the signal from the nodes does not fade in high extend. For large premisses, cluster based topology is preferred over mesh as in the latter higher energy is required at the nodes due to the fact that each node transmits its own readings and the readings of other nodes. Fig. 3 illustrates the proposed WSN-FITR for one floor in a commercial center. In this case there are several temperature clusters that are controlling a set of nodes. The decision regarding the temperature regulation is distributed among all controllers. There is one main controller that has wired connection to the server in order to retrieve/store information towards the database at the server.

3 Localization and Clustering in Indoor WSN

Many algorithms have been proposed for ZigBee based WSN localization. Most of them consider a WSN deployed in outdoor environment where GPS signals are available and a global map of the network can be easily achieved using well known techniques for localization. On the other side, indoor localization methods should consider different characteristics of the indoor surroundings where WSN is installed. Finding position of indoor WSN is more challenging since GPS signal is heavily attenuated by building structures such as walls and roofs and there is absence of line of sight to some satellites [6]. With only few exceptions, the distances between the nodes of the network are necessary to be known for accurate location prediction. Different techniques are used to obtain the distances:

- RSSI (Receive Signal Strength Indicator)
- ToA (Time of Arrival)
- AoA (Angle of Arrival)
- TDoA (Time Difference of Arrival)

The techniques based on RSSI are easier to implement and don't require additional hardware, as all standard wireless devices possess features for measuring this value. But finding the relationship between the signal strength attenuation and the transmission distance in indoor environments is not a trivial task [7] [8]. Additionally, many other characteristics of indoor environments have to be considered, like temperature and humidity variations, furniture rearrangements, presence of human beings, etc.

Indoor localization methods can be divided into two main categories [9]:

- deductive methods. They take into account only the physical properties of signal propagation. They require the positions of the access points, radio propagation model and map of the environment.
- inductive (fingerprinting) method. They require a previous training phase where the system learns the RSS in each location. This phase can be very time consuming. In the next (positioning) phase, different matching algorithm can be used in order to find the unknown location.

In [9] the authors present an algorithm that combines the advantages of both deductive and inductive methods. This hybrid method reduces the training phase without a loss of precision. In [10] several matching fingerprinting algorithms are investigated: the nearest neighbor (NN) algorithm, the K-weighted neatest neighbor (KWNN) algorithm and the probabilistic approach based on the kernel method. Through simulations it has been shown that KWNN algorithm has the best indoor positioning result.

Since localization is very crucial in our WSN-FITR, the algorithm should be selected very carefully in accordance with the characteristic of the environment where the network should be deployed. If there are many walls and obstacles in the environment, the deductive methods should be avoided because they estimate the position mathematically. When there are multiple access points and few walls in the environment, inductive methods are not necessary as the training phase can be very expensive.

After determining the location, sensor nodes in WSN can be geographically grouped into clusters. In each cluster one representative node (cluster-head) is chosen to coordinate member nodes. The main advantages of WSN clustering is not only to prolong the WSN lifetime, but also to establish collaboration between cluster members in order to provide data aggregation and more accurate reports about the region they sense. Many algorithms have been proposed in the literature for WSN clustering [11] [12].

4 Reductions of Data Transmissions

By reporting data measurement at each interval, WSN nodes consume a great deal of energy, which reduces its lifetime and creates sufficient communication overhead. Several techniques have been developed to overcome these problems i.e., to lower the communication overhead and to increase energy saving. Most of them consider reducing the number of radio transmissions.

Three main paradigms can achieve reduction of radio transmissions:

- *data compression;* where well known compression techniques are used to compress consecutive measurements. This approach is useful only if the WSN application doesnt require the data in real-time.
- *in-network processing;* when data are processed on their way to the sink. This method is usually performed when summarization functions or other queries are needed. It is appropriate only for mesh-based, cluster-based or hybrid-based network topologies.
- *data prediction;* when different prediction methods are used for predicting next sensor readings. Here, each node runs a filter (or a model) that estimates next sensor reading. The sink runs exactly the same models for each sensor in the network and makes the same predictions. This approach is known as Dual Prediction Scheme (DPS).

For the WSN-FITR system we want sensor measurements up-to-date, hence data compression is not an appropriate solution. Regarding network topology, we can choose among data prediction and in-network processing. If the network topology of WSN-FITR is star-based, we should apply data prediction methods. For different topologies we should consider in-network processing or combination of both.

In order to compare these techniques for data reduction, different algorithms were implemented in MatLab. For the evaluation, a set of experimental data from Intel Lab [17] was used. The 54 Mica2Dot sensors deployed in the laboratory were equipped with weather boards and measured temperature once every 31 seconds. The measurements were collected between February 28th and April 5th, 2004. We run the simulations for 50 different error margins eMax (ranging from 0.1° C to 5° C).

4.1 Data prediction using LMS-VSS

The most appropriate models (filters) for DPS are based on time-series forecasting: Moving Average (MA), Autoregressive (AR) [13], Autoregressive Moving Average (ARMA) [14], Least Mean Square (LMS) [15] and LMS with Variable Step Size (VSS-LMS) [16]. We implemented and evaluated LMS, LMS-VSS, second-order MA and ARMA. Fig. 4 shows the reduction gain for each of these algorithms simulated on two nodes form the Intel Berkeley Research Lab network [17]. The metric is the reduction of transmissions in percentage (Fig. 4, upper) and the difference between the predicted and the true value (Fig. 4, lower), i.e. mean square error (MSE).

4.2 In-network processing in WSN-FITR

For cluster, mesh and hybrid-based network topologies, each sensor readings should be retransmitted at least twice, so data reduction is expected to be smaller compared with star-based topology. But if WSN in deployed on vast region, the network can not be organized as star-based since radio signals are far from the sink and multi-hop routing is the only way for data to reach its destination.

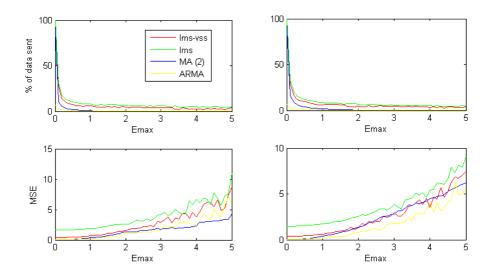


Fig. 4. Data reduction for different algorithms for Node 13 (left) and Node 49 (right).

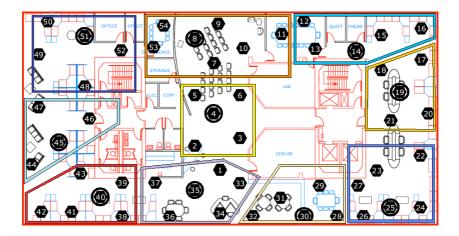


Fig. 5. A clustered view of the Intel Lab [17].

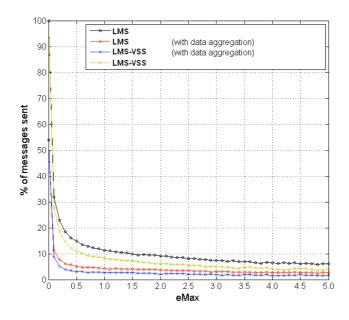


Fig. 6. Data reduction in the cluster containing the nodes: 7, 8, 9, 10, 11, 53 and 54.

In order to calculate the reduction in this case, we divided Intel network [17] into clusters (Fig. 5). Two algorithms were used for evaluation, LMS [15] and LMS-VSS [16]. The clustering parameter was geographic position, i.e., Euclidian distance. We assume that each sensor sends its reading to the cluster head, responsible for resending the reading to the sink. As a result, each reading is sent twice, except the readings taken at the cluster head. Additional reduction can be achieved if cluster head performs summarization function (Average, Minimum, Maximum, etc.) and forwards only the calculate aggregate to the sink.

Fig. 6 shows the reduction for cluster containing nodes: 7, 8, 9, 10, 11, 53 and 54. LMS-VSS shows an average gain of 5% compared to LMS algorithm. When cluster head performs data aggregation, the data reduction is far greater (97% reduction of the total messages sent for the given error margin of 0.5° C).

5 Conclusion

In this paper we propose WSN framework for indoor temperature regulation. We give an overview of the methods that can be used for nodes localization and clustering in ZigBee-based network. In order to reduce the energy consumption, we propose data reduction strategy based on dual prediction scheme that uses Least Mean Square filter as a prediction method. Through simulations on real world dataset we show that this filter is good predictor for temperature readings.

References

- Mittal, R.; Bhatia, M. P S, "Wireless sensor networks for monitoring the environmental activities," Computational Intelligence and Computing Research (ICCIC), 2010 IEEE International Conference on , pp.1,5, 28-29 Dec. 2010
- 2. ZigBee Alliance, (Accessed on 06/12/2012) http://www.zigbee.org/
- Chunlong Zhang; Min Zhang; Yongsheng Su; WeiLian Wang, "Smart home design based on ZigBee wireless sensor network," Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on , pp.463,466, 8-10 Aug. 2012
- Khan, A.A.; Mouftah, H.T., "Energy optimization and energy management of home via web services in smart grid," Electrical Power and Energy Conference (EPEC), 2012 IEEE , pp.14,19, 10-12 Oct. 2012
- Rossello-Busquet, A.; Soler, J.; Dittmann, L., "A Novel Home Energy Management System Architecture," Computer Modelling and Simulation (UKSim), 2011 UkSim 13th International Conference on, pp.387,392, April 1 2011
- Long Cheng, Chengdong Wu, Yunzhou Zhang, Hao Wu, Mengxin Li, and Carsten Maple, A Survey of Localization in Wireless Sensor Network, International Journal of Distributed Sensor Networks, vol. 2012, Article ID 962523, 12 pages, 2012.
- A. Awad, T. Frunzke, F. Dressler, Adaptive distance estimation and localization in WSN using RSSI measures, in Proc. 10th Euromicro Conf. Digital Syst. Des. Architectures, Methods Tools (DSD) 2007, pp. 471478.
- A. Faheem, R. Virrankoski, M. Elmusrati, Improving RSSI Based Distance Estimation for 802.15.4 Wireless Sensor Networks, in Conference Proceedings of the 2010 IEEE International Conference on Wireless Technology and Systems, Aug. 2010.
- 9. Lloret, J.; Tomas, J.; Garcia, M.; Canovas, A. A hybrid stochastic approach for self-location of wireless sensors in indoor environments. Sensors 2009, 9, 36953712.
- Jan, S.-S., Hsu, L.-T., Tsai, W.-M.: Development of an Indoor Location Based Service Test Bed and Geographic Information System with a Wireless Sensor Network. Sensor 10, 29572974 (2010)
- W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, An Application-Specific Protocol Architecture for Wireless Microsensor Networks, IEEE Trans. Wireless Commun., vol. 1, no. 4, Oct. 2002, pp. 66070.
- A.A. Abbasi and M. Younis, A survey on clustering algorithms for wireless sensor networks, Computer Communications, 30, 28262841, 2007.
- Y. L. Borgne , S. Santini , and G. Bontempi, Adaptive Model Selection for Time Series Prediction in Wireless Sensor Networks, Signal Processing, vol.87 n.12, pp. 3010-3020, December, 2007.
- J. Lu, F. Valois, M. Dohler, M.Y. Wu, Optimized data aggregation in WSNs using adaptive ARMA. International Conference on Sensor Technologies and Applications, SensorComm, Venice, Italy, July 2010. pp. 115120.
- S. Santini and K. Römer: An Adaptive Strategy for Quality-Based Data Reduction in Wireless Sensor Networks, Proceedings of the 3rd International Conference on Networked Sensing Systems (INSS 2006), pp. 29-36, Chicago, IL, USA. June 2006.
- B.R Stojkoska , D. Solev, D.Davcev, Variable step size LMS Algorithm for Data Prediction in wireless sensor networks, Sensor & Transducers Journal ,vol. 14-2 ,Special Issue, pp. 111-124, Feb 2012 .
- 17. Intel lab data. Web Page, (Accessed on 06/06/2011) http://db.lcs.mit.edu/labdata/labdata.html

Validation Study for Two-Stage Vessel Segmentation Algorithm

Teodoru Radu Popa¹, Mihai Mocanu¹

¹Faculty of Automation, Computers and Electronics, University of Craiova, RO-200440 Romania e-mail: tpopa@ software.ucv.ro

Abstract. A reduction of the size of the coronary artery openings, estimated by the visual inspection of the X-ray images of the transient radiocontrast distribution during conventional angiography, can be used to formulate predictions about clinical symtoms and dynamic reductions in coronary blood flowing. Even -with introduction of new noninvasive imaging modalities such as Cardiac CT and MRI, coronary angiography remains the gold standard for detecting ischmemic coronary disease. In this paper we describe an semiautomatic algorithm to segment and quantitative analyze the 2D coronary anatomy which offers great advantage from observer influences and bias, therefor minimizing significantly intraobserver and interobserver variability.Currently quantitative approaches to the angiographic evaluation of coronary anatomy are infrequently used on a routine daily basis in most angiographic laboratories because current devices don't offer an easy method to perform this task automatically and require an great effort from practitioners. The method involves combination of two well known segmentation techniques: fast marching and region growing by using the minimal geodesic path extracted from fastmarching as input for region growing.We validated the results by comparing using different comparison metrics(dice, jaccard) the segmentations performed by a physician with the output of our method. The results showed reasonable agreement(70%) between the gold standard and our method.

Keywords: vessel segmentation, region growing, fast marching, minimal geodesic path, STAPLE, coronary angiography

1 Introduction

A reduction of the size of the coronary artery openings, estimated by the visual inspection of the X-ray images of the transient radio contrast distribution during conventional angiography, can be used to formulate predictions about clinical symptoms and dynamic reductions in coronary blood flowing. Even with introduction of new noninvasive imaging modalities such as Cardiac CT and MRI, coronary angiography remains the gold standard for detecting ischemic coronary disease [2].Visual analysis generally leads to overestimation of severe stenoses, and to underestimation of more modest degrees of luminal narrowing [1].Currently quantitative approaches to the angiographic evaluation of coronary anatomy are infrequently used on a routine daily basis in most angiographic laboratories because

current devices don't offer easy solutions to perform this task automatically and require a great effort from practitioners.

In this paper we present a novel two stages segmentation approach with the goal to improve upon the several factors that affect manual segmentation (length, shape and eccentricity).

This paper is structured in two parts. In first part we describe an semi-automatic algorithm to segment and analyze the 2D coronary anatomy which offers great advantage from observer influences and bias, there for minimizing significantly intraobserver and inter-observer variability.

In second part we validate our algorithm using metrics used in the literature.

2 Two stage segmentation algorithm

The hybrid method involves combination of two well-known segmentation techniques: fast marching and region growing by using the minimal geodesic path extracted from fast marching as input for region growing.

The Proposed Overall Architecture is described in the Figure 1.

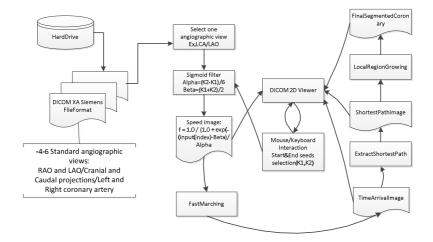


Fig. 1. Proposed Segmentation Architecture

First we perform fast marching segmentation on the input image. The output of this algorithm it is used to extract the geodesic minimal path. The minimal path contains a list of points which are used as seeds point for a local region-growing segmentation.

Fast Marching and LevelSet Segmentation Methods Overview

In the next section we will briefly describe the general level set methods and we will detail our implementation of one of the level set methods: the fast marching.

LevelSet Segmentation and Fast Marching are numerical methods that can follow evolution of an interface .These methods can be used to segment an object in an image or to find the minimal geodesic path between two points. Also levelset methods have a wide range of application in fluid dynamics, combustion and image processing. Both were developed by James A Sethian as numerical methods to solve viscosity interfaces based on Eikonal equations: [3]

$F(x)|\nabla T(x)| = 1$

This equation describes the evolution of a closed curve as a function of time T with a speed F(x) in the normal direction at a fixed point x. The speed function is given and the moment of time the contour intersects the point x is obtained by solving this Eikonal equation.

The order in which the values of this image grid are produced by finit differences approximations is closely correlated with Dijkstra method. This method is a technique for depth searching for calculating the shortest path inside a grid. The algorithm maintains a history of speed propagation along the grid points, by walking along the grid points to reach every grid point. Fast Marching method exploits a similar concept of finite difference approximations for calculating the partial differential equation rather than using the actual grid points finite values.

The Fast Marching method uses as a speed function the normalized magnitude of the gradient of the image, in which by gradient we mean how much the neighbor pixels differ in intensity value and normalization of the gradient translate the values inside the interval 0-1.

Fast Marching Algorithm

In the following section we will describe the algorithm's implementation.

We need to solve an equation by computing the arrival time T at each point (x, y)based on its neighbors.

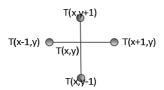


Fig. 2. FastMarching 4 neigbour points for T(x,y)

The quadratic equation as described in [4] based on upwind scheme is: $aT_{x,v}^{2} + bT_{x,v} + c = 0$ (1)Where coefficients a,b,c are defined as:

$$a = \left(\frac{1}{\Delta^{2}x} + \frac{1}{\Delta^{2}y}\right)$$

$$b = \left(\frac{\min(T_{x-1,y}, T_{x+1,y})}{C^{2}x} + \frac{\min(T_{x,y-1}, T_{x,y+1})}{C^{2}y}\right)$$

$$c = \left(\frac{\min(T_{x-1,y}, T_{x+1,y})^{2}}{\Delta^{2}x} + \frac{\min(T_{x,y-1}, T_{x,y+1})^{2}}{\Delta^{2}y}\right) - \frac{1}{F_{x,y}^{2}} (2)$$

By solving this equation (2) we obtain two solutions and we pick the largest value because it satisfies the Fast Marching rules. The type of points for min() coefficients must be blacks otherwise we will use zeros instead.

Following are described the common steps used in implementation of this algorithm.

Initialization

1. We set T(X0)=0 and we mark all the points in the starting band as black.

2. We set the other point as T(X0)=1 and we mark these points with green. Iteration step

1. Green points adjacent to black will become red.

2. Red points are updated and their time T is estimated based on arrival times of the neighbor black points by calculating the Eikonal equation described above.

3. Red points with the smallest T become black points.

4. The process is reiterated until there are no longer red points available.

One observation is that a point is updated only using the values of the neighbor points (4 neighbor points close to it) with the smallest T value, therefore the new value of T is bigger of the point's value after the update.

Fast Marching minimal path extraction is a semi-autonomous segmentation method: the user is required to provide start and end points.

In Figure 3 the example of coronary artery segmentation is shown. The coronary data has good contrast and dimensions 512x512 and is of XA DICOM format. The coronary has been partially segmented starting with 2 seeds (start and end seed) as input data.

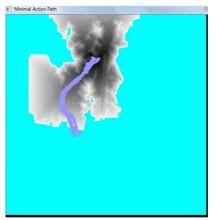
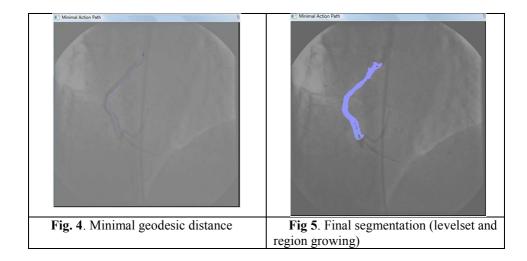


Fig. 3. Time Arrival Image



Extracting geodesic minimal path

After performing fast marching segmentation we perform the next step: extracting an optimal path from a starting position to an end point. To perform this step we iteratively backtrack from the end point towards the starting point by using the minimal values of arrival time map at each iteration.

In Figure 4 is shown the geodesic minimal path colored as a blue curve. Red point is the starting point and green point represents the end point

Region Growing Segmentation Step

The next step of the hybrid segmentation is represented by the local region growing segmentation. Region-growing approaches exploit the important fact that pixels which are close together have similar gray values. Through performing a restricted and local segmentation we can alleviate the leaking effects of the global fast marching method. The local segmentation ensures that only pixels similar in value with the local seed will be used instead of the original starting seed.

And by applying region growing segmentation we ensure that the fine details will be preserved compared to local level set which has the tendency to smooth the segmentation result.

Region growing algorithm involves the next steps:

Start with a single pixel (seed) and add new pixels slowly. Choose the seed pixel. In our case we use every point included in the minimal path

Check the neighboring pixels and add them to the region if they are similar to the seed

Repeat step 2 for each of the newly added pixels; stop if no more pixels can be added or a certain segmentation diameter is reached. We define a constant circle segmentation zone around each seed point by providing a segmentation diameter.

One of the advantages of adding this region growing step is the no single region is allowed to completely dominate the overall segmentation. This step can be further improved by allowing all of regions to grow at the same time. Similar regions will gradually combine into expanding regions. Control of these methods may be quite complicated but in our case we developed a simple addition method by adding the new segmentations to the old ones. Also by doing so it's easy and efficient to implement on parallel computers.

In Figure 5. We showed the results of the final segmentation.

3 Validation Study

To compare the results of our segmentation method we can apply several measures from literature. In this paper we measure binary segmentations for accuracy, which simplifies the complexity of the measures algorithms used.

There are several measures in literature. One of the measures that is currently used is Global consistency error(GCE). This measure was introduced by [5] in his PHD dissertation thesis and it measures the extent to which one segmentation can be viewed as a refinement of the other. Segmentations which are related in this manner are considered to be consistent, since they could represent the same natural image segmented at different scales.

The Global consistency error measure allows comprehensive experiments to measure the performance of an investigated algorithm in terms of visual evaluation and a variety of quantitative indices for image segmentation.

The Global consistency error measure needs a ground truth for comparison. Unfortunately, GCE only vaguely describes the ground truth concept. Also, despite being popular and robust, it is better suited to multi-segmentation algorithm.

As we don't consider that there is only one validation measure to compare the binary segmentation results, we used a set of measures based on current literature. These validation measures are implemented as a validation toolkit in the open source library ITK (www.itk.org).

We used the following validation measures: Dice index, Jaccard index, Volume Overlap, Volume Similarity, False negative and false positive measures.

To tackle the robust ground truth construction we used a different algorithm: STAPLE. This filter, named the Simultaneous Truth and Performance Level Estimation algorithm, generates ground truth volumes from a set of binary expert segmentations.

STAPLE Algorithm

The ground truth segmentation was estimated from a set of manual segmentations performed by one experienced cardiology resident. The result generated by the semiautomatic segmentation was evaluated by comparison with the estimated ground truth. We estimated ground truth from the manually segmented data by binary thresholding of the result of the STAPLE algorithm with a constant value (95% probability) [10]. To estimate the ground truth from multiple experts we applied the expectation maximization method described by Warfield.. This method uses binary labeled images and also estimates the performance parameters of each expert in terms of sensitivity and specificity. This method was also used to determine the most accurate and the least accurate value obtained from manual segmentation by each of the experts, relying on the final values of sensitivity and specificity.

The expectation maximization has four steps [11]:

1) Initialization- the hypothesis is initialized with the average segmentation value of all radiologists.

2) Estimation– the estimation of the likelihood of P(D,T | p,q) is based on the current hypothesis (p',q') and the observed data D. We use the expected value $E(\ln P(D,T | p,q))$ because the full data (D,T) is considered here to be a random variable.

3) Maximization– the hypothesis (p,q) is replaced by the new hypothesis (p',q'), that maximizes the estimate of the likelihood P(D,T | p,q).

4) Convergence– this is determined by calculating the rate of change of the sum of the true segmentation probability.

Next figure 6 describe these steps:



Fig. 6. STAPLE General Expectation Maximization Framework.

4 Results

One medical resident segmented the gold standard images, 5 datasets for each test. For each set of 5 images we extracted an estimated ground truth used later for validation.

In table 1 and 2 is shown the output of STAPLE algorithm: sensitivity, specificity and number of iterations.

The STAPLE algorithm also outputs an estimated ground truth.

Table 1. Ground truth estimation results using STAPLE algorithm for 5 human
experts and the first 3 test datasets.

	DataSet1	Set1 DataS		2	DataSet3	3
	Sens*	Speci*	Sens*	Speci*	Sens*	Speci*
Expert1	0.85759	0.99869	0. 6151	0.999656	0.8512	0.9932
Expert2	0.9728	0. 9994	0.9531	1	0.9012	0.9919
Expert3	0.95413	0.	0.99713	0.999413	0.9715	0.9981

		99987				
Expert4	0.90055	0.99985	0.84158	0.99866	0.9937	0.9988
Expert5	0.99985	0. 9999	0.9521	0.9991	0.9034	0.9982
Mean	0. 9279	0. 9995	0.851	0.9994	0.9247	0.99604
Nr itr	5		16		8	

* Sens=Sensitivity(p);Speci=Specificity(q)

Table 2. Ground truth estimation results using STAPLE algorithm for 5 human
experts and the last 2 test datasets.

	DataSet4		DataSet5	
	Sens*	Speci*	Sens*	Speci*
Expert1	0.701	0.9923	0.63984	0.99956
Expert2	0.8532	0.9971	0.984132	0.999881
Expert3	0.9511	0.993	0.948303	1
Expert4	0.9531	1	0.72232	0.98122
Expert5	0.90122	0.99232	0.95543	0.99232
Mean	0.871924	0.994944	0.850006	0.994596
Number	13		29	
iterations				

* Sens=Sensitivity(p);Speci=Specificity(q)

The estimated ground truth images were compared with the results obtained by the two-stage hybrid segmentation method. The comparisons values from segmentation are shown in the next table(table 3).

We used the ITK's(www.itk.org) well known comparison metrics to perform this task.

Study Number	Total	Union (jaccard)	Mean (dice)
1	0. 748165	0. 786699	0. 648396
2	0.748165	0.786699	0.648396
3	0.743673	0.780053	0.639415
4	0.910645	0.915369	0.843945
5	0.839001	0.864591	0.76148
Study Number	Volume	False negative	False positive
	similarity.		
1	0. 207665	0. 251835	0. 582288
2	0. 207665	0. 251835	0. 582288
3	0.186523	0.256327	0.607801
4	-0.492935	0.0893547	0.615523
5	0.037004	0.160999	0.460017

 Table 3. Validation results

From the Table 1,2 and 3 we can conclude that the overall match was around 70% for all 5 datasets. Despite a lower agreement between the gold standard and segmentation results we consider the method useful for analysis of coronary stenosis

in some cases and shows good agreement between the 2 segmentations. One explanation for the disagreement is that in a clinical setting the physicians usually do not segment the dataset to calculate the degree of stenosis and rather estimate it visually. Also the method is useful because the required accuracy of stenosis is permissive, allowing +10% error without modifying the outcome of the diagnosis.-

2 Conclusion

In this paper we described and implemented an algorithm for segmenting coronary arteries on X-Ray 2D Angiographic images. Despite there are many algorithms for segmentation and analysis of the coronary angiographic images, there is not a common practice and a trivial task in cardiology to perform these kinds of procedures.

The algorithm implemented is based on a hybrid approach and offers the benefits of both algorithms on which are based: the robustness of fast marching methods and the precision and fine details of region growing.

From the validation study we can conclude that the hybrid approach offered reasonable results compared to a certified assistant physician, but these are only preliminary results. We tested only on 5 datasets which the method performed well.

Further we will test this method on a larger database. Also this method is slower that either region growing or level set alone because it combines the two with a penalty in execution time.

Acknowledgement

This work was supported by the strategic grant POSDRU/89/1.5/S/61968, co-financed by the European Social Fund within the Sectorial Operational Program Human Resources Development 2007-2013.

References

- Fleming RM, Kirkeeide RL, Smalling RW, Gould KL. Patterns in visual interpretation of coronary arteriograms as detected by quantitative coronary arteriography. J Am Coll Cardiol; 18:945.(1991)
- 2. White CW, Wright CB, Doty DB, et al. Does visual interpretation of the coronary arteriogram predict the physiologic importance of a coronary stenosis? N Engl J Med; 310:819.(1984)
- 3. J. Sethian. : ,Level Sets Methods and Fast Marching Methods. Cambridge University Press, 2nd edition, (1999).
- A. Farag and M. Sabry Hassouna: Theoretical Foundations of Tracking Monotonically Advancing Fronts Using Fast Marching Level Set Method, Technical Report, Louisville, Feb (2005)

- Martin, D., Folks, C., Tal, D., Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proc. 8th Int'l Conf. Computer Vision. Volume 2. pag.416-423 (2001)
- DeRouen TA, Murray JA, Owen W. Variability in the analysis of coronary arteriograms. Circulation ; 55:324.(1977)
- 7. Goldberg RK, Kleiman NS, Minor ST, et al. Comparison of quantitative coronary angiography to visual estimates of lesion severity pre and post PTCA. Am Heart J; 119:178.(1990)
- 8. Zir LM. Observer variability in coronary angiography. Int J Cardiol; 3:171.(1983)
- 9. Shub C, Vlietstra RE, Smith HC, et al. The unpredictable progression of symptomatic coronary artery disease: a serial clinical-angiographic analysis. Mayo Clin Proc ; 56:155.(1981)
- S K Warfield, K H Zou, and W M Wells. :Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. IEEE Trans Med Imaging 23 (7):903-21.(2004)
- 11. S K Warfield, K H Zou, and W M Wells. :Validation of image segmentation by estimating rater bias and variance. Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv 9 (Pt 2):839-47.(2006)
- 13. Epstein, SE, Cannon, RO III, Watson, RM, et al. Dynamic coronary obstruction as a cause of angina pectoris: Implications regarding therapy. Am J Cardiol; 55:61B(1985).
- 14.Feldman RL, Nichols WW, Pepine CJ, Conti CR. Hemodynamic significance of the length of a coronary arterial narrowing. Am J Cardiol ; 41:865(1978).
- 15.Feldman RL, Nichols WW, Pepine CJ, et al., The Coronary hemodynamics of left main and branch coronary stenoses. The effects of reduction in stenosis diameter, stenosis length, and number of stenoses. J Thorac Cardiovasc Surg ; 77:377.(1979)

Society and Human Development Psychology, Politcs, Sociology and Education

The relationship between dysfunctional patterns of families' response to the illness and symptom severity in adolescent patients with anorexia nervosa at illness onset: A gender-specific approach

Dimitra Anastasiadou¹, Ana R. Sepulveda¹ and Montserrat Graell¹

¹ Department of Biological and Health Psychology, Faculty of Psychology, Autonomous University of Madrid

²Section of Eating Disorders, Child and Adolescence Psychiatry Service, Niño Jesus University Hospital, Madrid, Spain.

Email: dimi.anastasiadou@gmail.com

Abstract. The objective of this study is to take a snapshot of the family environment and identify, integrating a gender perspective, factors related to dysfunctional patterns of family's response to the illness that may be associated with the severity of Eating Disorder (ED) symptoms at Anorexia Nervosa (AN) onset in adolescents. A total of 48 mothers and 45 fathers of 50 patients aged from 12 to 18 years, with a mean duration of illness of 15.74 months (SD = 12.58) all diagnosed with AN according to DSM-IV, were recruited from an ED Unit in Madrid (Spain) on the first day they were referred for treatment. Caregivers completed self-report assessments related to their anxiety and depression levels (HADS), quality of life (SF-36), caregiving experience (ECI), ED symptom accommodation (AESED), Expressed Emotion (FQ) and coping mechanisms (COPE-24) and patients went through clinical interviews examining ED symptoms (EDE-12) and comorbidity with other mental disorders (K-SADS). Mothers showed higher levels of anxiety and emotional over-involvement and exhibited a poorer quality of life and a more negative caregiving experience compared to fathers. In turn, fathers showed more adaptive coping mechanisms to handle the illness of their daughters/sons compared to mothers. Different associations between psychological variables related to family's response to the illness and ED attitudes and behaviors in patients were found for fathers and mothers. Interventions that help parents to cope with their caregiving role should target behavioral, cognitive and emotional responses, and be gender-specific for mothers and fathers, thereby possibly improving the outcome of AN in patients.

Keywords: caregiving experience, family functioning, gender, anorexia nervosa

1 Introduction

Having a child with a mental disorder may lead to a worsening of the caregiving experience in families, considering that families are living together with the patient, spending many hours in contact with the patient and also having to assume the caregiver's role, all in addition to their essential role in their child's development and education.

Caring for a relative with an Eating Disorder (ED) is associated with significant levels of psychological distress and burden. Moreover, families of persons suffering from an ED report physical and mental health deterioration, poor quality of life as well as interpersonal difficulties, including social stigma and isolation.

Gender differences related to the caregiving experience and the family's coping responses to the illness should be examined in order to gain a more comprehensive picture about risk and protective factors in the family environment.

Several models have been developed to better explain the caregiving experience and understand how to intervene on the dysfunctional patters of caregiving in order to improve illness outcomes and reduce caregivers' levels of distress. However, none of these models approaches the caregiving experience from a gender perspective.

2 Literature review

Families' response to the illness can be a maintaining factor in EDs and a predictor of poor outcome in patients with Anorexia Nervosa (AN) (Schmidt and Treasure, 2006). More specifically, ED symptoms and behaviors may trigger distress in carers, which in turn may lead carers to adopt maladaptive coping efforts through symptom accommodation or high levels of Expressed Emotion (EE), leading to an increased caregiving burden, physical and psychological health problems and significant emotional distress (Dimitropoulos, Carter, Schachter, & Woodside, 2008; Kyriacou, Treasure, & Schmidt, 2008b; Sepulveda et al., 2009). Finally, the above-mentioned family factors may lead to a worsening of ED symptoms and behaviors in patients with AN.

In regard to adolescents with AN, families can play an active role in the treatment process by helping with their child's eating attitudes and behaviors and weight restoration (Lock et al., 2010).

Finally, it is is important to note that there is an underpresentation of fathers in clinical research, with them being the "neglected" group, thereby perpetuating the

stereotype of an overinvolved mother and a peripheral father (Cook-Darzens et al., 2005).

The first objective of this study is to take a snapshot of family environment by describing the caregiving experience of relatives of people with an ED in terms of physical and mental health, burden, quality of life, EE and coping strategies, integrating a gender perspective. The second objective is to identify dysfunctional patterns of family response to the illness that may be associated with the severity of ED symptoms in mothers and fathers of adolescent patients with AN at onset of their illness.

3 Data and Methodology

Sample

48 mothers with a mean age of 44.9 years (SD = 4.6) and 45 fathers with a mean age of 47.5 years (SD = 4.1) of 50 patients aged from 12 to 18 years, with a mean duration of illness of 15.7 months (SD = 12.6) all diagnosed with AN, were recruited in an ED Unit in Madrid, Spain, the first day they were referred for treatment.

Statistical analysis

To examine the differences between fathers and mothers in psychological variables, a non-parametric Mann-Whitney U test was performed. Pearson's correlation coefficients between psychological variables of fathers and mothers and patients' symptom severity index were performed. Statistical significance was set at p<0.05.

Instruments

- Patients
 - Clinical Interviews
 - Eating Disorder Examination (EDE-12; Fairburn & Cooper, 1993).

K-SADS- PL (Kaufman, Birmaher, Brent, Rao, & Ryan, 1996).

• Carers

- Accommodation and Enabling Scale for Eating Disorders (AESED; Sepulveda, Kyriacou & Treasure, 2009) consisting of five factors: 1) Avoidance and Modifying Routine, 2) Reassurance Seeking, 3) Meal Ritual, 4) Control of Family and 5) Turning a Blind Eye.
- COPE-24 (Sica, Novara, Dorz, & Sanavio, 1997; Spanish validation in student sample by Crespo and Cruzado, 1997) which consists of 14 dimensions as follows: self-distraction, active coping, denial, substance use, social support, behavioral disengagement, venting, positive reframing, planning, humor, acceptance and religion.
- Experience of Caregiving Inventory (ECI; Szmukler et al., 1996; Spanish validation by Sepulveda et al., *in press*) which consists of two subscales: 1)
 Negative caregiving experience and 2) Positive caregiving experience.
- **Family Questionnaire** (Wiedemann et al, 2002; Spanish validation by Sepulveda et al., *in preparation*) which consists of two subscales: 1) Emotional Overinvolvement and 2) Criticism.
- Hamilton Anxiety and Depression Scale (HADS; Sigmond & Snaith, 1983; Spanish validation by Herrero et al., 2003) which consists of two subscales: 1) Anxiety and 2) Depression.
- SF-36 Health Survey (Ware and Sherbourne, 1992; Spanish validation by Ware et al., 1998) with 1) Physical health component and 2) Mental health component.

4. Empirical Analysis: Preliminary results

	Ν	Mean	Standard Deviation	Range
Age (years)	50	14.78	1.79	12-18
Illness duration (months)	50	15.74	12.58	2-48
Age of illness onset	50	13.80	1.68	11-17
BMI	50	18.14	1.82	14.86- 23.20
	Ν	%		
Treatment type				
Outpatient (Ambulatory care)	10	20		
Outpatient ("Home hospitalization")	20	40		
Outpatient (Day hospital)	3	6		
Inpatient	17	34		
Sex				
Female	49	98		
Male	1	2		
Nationality				
Spanish	45	90		
Other	5	10		

 Table 1. Sociodemographic and clinical data of patients.

 Table 2. Sociodemographic and clinical data of families.

	Total		Mothers		Fathers		
	Ν	Mean (SD)	Ν	Mean (SD)	Ν	Mean (SD)	t; p- value
Age (years)	93	46.13(4.55)	48	44.88(4. 61)	45	47.47 (4.13)	2.849; 0.01
	Ν	%	Ν	%	Ν	%	χ2;p-val

							ие
Sex							
Male	45	48.4					2.860
Female	48	51.6					0.01
Educational level							
School/Secondary level	22	23.7	10	20.8	12	26.7	0.438;
Degree/Diploma	71	76.3	38	79.2	33	73.3	0.51
Marital status							
Married/Living together	78	83.9	40	83.3	38	84.4	0.021;
Single/Divorced/Widowed	15	16.1	8	16.7	7	15.6	0.88
Labour situation					_		
Full time/Part time	72	79.1	35	72.9	39	86.7	2.705
Unemployed/Retired	19	20.9	13	27.1	6	13.3	0.10
Living with the patient							
Yes	84	90.3	45	93.8	39	86.7	1.333;
No	9	9.7	3	6.3	6	13.3	0.248
Contact hours							
$\geq 21h/week$	72	77.4	41	85.4	31	68.9	3.629;
<21h/week	21	22.6	7	14.6	14	31.1	0.05
History of eating- and weight-related problems							
Yes			10	20.8	5	11.1	1.623;
No			38	79.2	40	88.9	0.20
Psychiatric history							
Yes			13	27.1	6	13.3	2.701
No			35	72.9	39	86.7	0.10

Regarding preliminary results of the present study, mothers showed higher levels of anxiety and emotional overinvolvement and a poorer quality of life (physical and mental health component) compared to fathers. Moreover, they exhibited a more negative caregiving experience and at the same time they appreciated to a greater degree the positive aspects of their relationship with their daughters/sons, compared to fathers. In turn, fathers used more adaptive coping mechanisms to cope with the ED of their daughters/sons, using humor and planning as coping mechanisms and also seeking social support.

In addition, ED attitudes and behaviors were positively associated with emotional overinvolvement, criticism, a negative caregiving experience, anxiety and depression scores, denial and use of religion as coping strategies, in mothers. In fathers, symptom severity index was associated with criticism, anxiety, symptom accommodation and denial as a coping mechanism.

5 Conclusions

In the Doctoral Thesis of Raenker (2011) and in accordance with our findings, mothers expressed more feelings of distress and burden than fathers, whereas fathers expressed more self-efficacy as caregivers compared to mothers using more adaptive coping strategies. Moreover, the findings of the present study confirm the findings of Kyriacou et al. (2008a) and Whitney et al. (2005) in which mothers of patients with AN expressed more emotional overinvolvement, anxiety and depression than fathers.

In studies with other groups of patients, such as families of patients with dementia, Alzheimer, other forms of disabilities, or caregivers of elderly people, and similarly to our findings, mothers exhibited more negative scores in quality of life (Collins & Jones, 1997; Rose-Rego et al., 1998), they also showed more caregiving burden (Collins & Jones, 1995; Fredriksen, 1996; Kramer & Kipnis, 1995) and a more depressive symptomatology (Barusch & Spaid, 1989; Lutzky & Knight, 1994; Pruchno & Resch, 1989; Rose-Rego, Strauss, & Smith, 1998) compared to fathers.

Both mothers' and fathers' anxiety levels, criticism and denial of the illness were associated with severity of ED attitudes and behaviors. Kluck (2008) also found an association between family dysfunction and increased disordered eating in a nonclinical sample of college women. An important additional finding observed in the present study that has not been previously reported, was that mothers' emotional and cognitive responses to the illness were associated with symptom severity in patients whereas in fathers, behavioral responses centered on unhelpful enabling and accommodating behaviors were associated with symptom severity.

A possible explanation for the gender-related differences is that fathers are often excluded from treatment decisions and they express uncertainty in regards to symptom management and lack of control. As a consequence, their coping response to the illness may be to avoid confrontations, to deny the illness and to get involved in unhelpful reassuring and enabling behaviors. On the other hand, mothers are usually primarily responsible for providing formal or informal care for their sons/daughters, spending more hours in daily contact with them and experiencing higher levels of burden. As a consequence, they become more anxious and emotionally overinvolved and they experience a poorer quality of life than fathers.

Another possible explanation related to the differences between the coping strategies of fathers and mothers in their coping strategies is that, since fathers usually play a more peripheral role caring for their children they may adopt more adaptive coping strategies by distancing themselves emotionally from the problem, using active coping responses such as humor, planning and seeking out social support.

Considering that the impact on carer's distress is different for fathers and mothers of adolescents with AN (Kyriacou et al., 2008a), interventions that help parents to cope with their caregiving role should target behavioral, cognitive and emotional responses adapted to each parenting style and perception of family functioning (Cook-Darzens, Doyen, Falissard, & Mouren, 2005; Treasure et al., 2007, 2009) taking into account gender differences.

References

Cook-Darzens, S., Doyen, C., Falissard, B., & Mouren, M. (2005). Self-perceived Family Functioning in 40 French Families of Anorexic Adolescents: Implications for Therapy. *European Eating Disorders Review*, 13, 223-236.

Kluck., A. (2008). Family factors in the development of disordered eating: Integrating dynamic and behavioral explanations. *Eating Behaviors*, 9, 471-483.

Raenker, S., Hibbs, R., Goddard, E., Naumann, U., Arcelus, J., Ayton, A., et al. (2012). Caregiving and Coping in Carers of People with Anorexia Nervosa Admitted for Intensive Hospital Care. *International Journal of Eating Disorders*, 00, 000-000.

Schmidt, U., & Treasure, J. (2006). Anorexia nervosa: Valued and visible. A cognitive- interpersonal maintenance model and its implications for research and practice. *The British Journal of Clinical Psychology*, 45, 1–25.

Sepúlveda, A. R., Graell, M., Berbel, E., Anastasiadou, D., Botella, J., Carrobles, J. A., & Porandé, M. (2012). Factors associated with emotional well-being in primary and secondary caregivers of patients with eating disorders. *European Eating Disorders Review*, 20, 1, e78-84. Whitney, J., Murray, J., Gavan, K., Todd, G., Whitaker, W., & Treasure, J. (2005). Experience of caring for someone with anorexia nervosa: qualitative study. *The British Journal of Psychiatry*, 187, 444-449.

Reading Literacy in the Educational Process from Primary School Teachers' Viewpoint

Mgr. Bc. Jana Drábová, Mgr. Tereza Gajďoková, Doc. PhDr. Vlasta Řeřichová, CSc.

Pedagogical Faculty (Department of Czech Language and Literature), Palacky University Olomouc

jana.drabova@centrum.cz

Abstract. The reading literacy is a man's attribute that develops over the entire course of life equipping the individual with knowledge, skills, capabilities, attitudes and values necessary for using all sorts of texts in various individual and social contexts. At present the current level of the reading literacy is considered to be rather poor, so it becomes one of the most pressing problems in the field of education all over Europe. This article provides information on the results of the research focused on the chosen topical issues of the reading literacy and their impact upon the educational process from the primary school teachers' viewpoint. Therefore, the empirical investigation has focused on primary school teachers. The investigation has been carried out in primary schools chosen by means of the simple random sampling method within the territory of the Czech Republic. The research sample has included approximately 400 primary school teachers. Thus, we have obtained data on the reading literacy level from the primary school teachers' point of view depending on the objective and subjective assumptions of its development.

The pedagogical research of the reading literacy has been based upon quantitative as well as qualitative research methods. The main research tool has been the non-standardized questionnaire method. The author of the article presents a questionnaire providing continuous results of the questionnaire investigation, which represents a part of the solution of the research project Reading Literacy from the Primary School Teachers' Viewpoint. One of the major qualitative methods has been observation. We have used the observation method in the lessons of literature education, Natural Science and Geography. Our aim has been to find out what strategies the teachers apply, and how often they use them, in order to develop the reading literacy. We have also been interested in pupils' skills and activities contributing to the development of their reading literacy.

Keywords: reading literacy, topical issues of reading literacy from the teachers' viewpoint, primary school teachers.

1 Introduction

In this article we present the research results up to now concerning some of the topical issues of the reading literacy and their impact upon the teaching-learning process from the primary school teachers' viewpoint. The article includes a questionnaire which was applied for the purpose of researching into the reading literacy, and which is a part of the solution of the research project within the Ministry of Education, Youth and Sports – Reading Literacy from the Primary School Teachers' Viewpoint. At the beginning we introduce the research project along with the executed investigation of the reading literacy topical issues. Then we try to explain the current situation in the field, and last but not least, we deal with pupils' activities that contribute to the development of the reading literacy along with the level of their knowledge.

The reading literacy, as a prerequisite of continuous education and successful integration of the individual into society, becomes the focal point of the curriculum in all developed countries.

The current situation of the reading literacy has been a subject of several international comparative researches (e.g. PIRLS, PISA, etc.). Teaching reading traditionally belongs to the primary education priorities in Czech schools. Nevertheless, what we know from the above researches is that there are more pupils with an unsatisfactory level of reading competences, which means that they might have difficulties in their further studies and also in everyday life. For one third of Czech pupils reading is one of their favorite activities, but on the other hand, one third of Czech pupils consider reading to be a waste of time (EACEA P9 EURYDICE. (2011) Teaching Reading in Europe: Contexts, Policies and Practices. Brussels: Education, Audiovisual and Culture Executive Agency).

For the reasons mentioned above our research focuses on topical issues of the reading literacy from the primary school teachers' viewpoint so that we can learn of the attitudes from the key representatives of the teaching-learning process, who substantially influence the current level of the reading literacy. Then we can draw adequate and applicable conclusions.

We try to find out what primary school teachers think about the reading literacy current level and how they view the possibilities of further development in this area. From the pedagogical point of view we pay attention to pupils' attitudes towards reading and their influence upon the reading literacy development, we also focus on the reading genre preferences in the lessons in the reading literacy development. One of the objectives is to confront the teachers' attitudes with those of pupils and with the possibilities of school literature education to motivate pupils to read for a lifetime.

2 Literature review

Reading literacy is becoming a very important issue worldwide. Reading literacy is seen as a functional part of literacy and at the same time it is considered to be means of achieving the literacy. Reading does not mean only decoding, but also includes active work with a text, critical judging and represents one of the most important means of further educational development. Due to its exceptional position reading literacy has become a part of many important international comparison researches, such as RLS, PISA, PIRLS.

In the Czech Republic a wide range of institutions and experts on the field of educational system has been dealing with reading literacy. To the most significant ones belongs Ministry of education, Czech school inspection and Pedagogical research institute, Institute for education information. In the Czech Republic the authors interested in Reading Literacy are mainly doc. PhDr. Ivana Gejgušová, Ph.D. (2012), PhDr. Radana Metelková Svobodová, Ph.D. (2011), Mgr. Jana Hyplová, Ph. D. (2010, 2011), Mgr. Veronika Najvarová, Ph. D. (2007, 2008, 2010), prof. PhDr. Milada Rabušicová, dr. (2002), Mgr. Petr Najvar, Ph.D. (2007). In other europian countries the authors interested in Reading Literacy are Prof. PhDr. Peter Gavora, CSc. (2008), Krashen, S. D. (2004), Temple, CH.; Ogle, D.; Crawford,A. Freppon, P. (2005), Barry, A. L. (2002), Kucer, S. B. (2005).

The first international research in reading literacy was in Czech schools processed in 1995.It was an international reading literacy research RLS. This research was organized by the International Association for Evaluation of Education Achievement (IEA). The Czech Republic has been its member since 1991. This research represented the first occasion to compare Czech pupils' knowledge and skills with students from other countries. Czech students achieved average results.

Lately two international researches, PIRLS (Progress in International Reading Literacy) organized by IEA and OECD/PISA (International Programme for Student Assessment), have been executed. PIRLS focusing on primary school pupils is especially interesting to us. PIRLS research was in Czech schools executed in years 2001 and 2011 and focused on the fourth class pupils. This research also included a questionnaire analysing the family background and school environment as important factors influencing reading literacy development.

Coming from the 2011 PIRLS results we can judge the reading literacy situation from pupils' as well as teachers' viewpoint. In total 8% of pupils achieved very high level reading literacy skills, 42% achieved high level skills, 37% achieved mediate reading literacy skills and 11% had low level skills. Less than 2% of pupils' skills were below low level. The Czech Republic scored over median.

Teaching methods have a big impact on reading literacy development. The most essential data collected from PIRLS (2011) concern the activities used for teaching reading, reading literacy development and the teaching materials used. In comparison to the other countries, Czech teachers do not use the whole scale of activities. Czech teachers prefer reading aloud (80%) and quiet reading (45%). To develop reading literacy skills teachers also use information looking up (99%). The least popular activity is text judging from authors' viewpoint (30%). The most common teaching materials are textbooks, used by 90% of teachers.

3 Data and methodology

Reading literacy is seen as a set of knowledge and skills of an individual, which enables him to work with written texts occuring in everyday life. It is not only a skill to read a text and to understand it, but it is also connected with a skill to look up, analyze and compare information and to reproduce a content of a text (METELKOVÁ – SVOBODOVÁ, R., HYPLOVÁ, J. (2011) Strategie rozvíjení čtenářské gramotnosti v čítankách 1.stupně ZŠ. Ostravská univerzita v Ostravě, Ostrava.)

Reading literacy is a skill of understanding and using the forms of written language required by the society. Young readers can derive the meanings from a wide range of texts. They read to learn, to become a part of the society and to have fun (METELKOVÁ – SVOBODOVÁ, R., HYPLOVÁ, J. (2011) Strategie rozvíjení čtenářské gramotnosti v čítankách 1.stupně ZŠ. Ostravská univerzita v Ostravě, Ostrava).

Several levels mingle in the reading literacy:

• A positive attitude to reading – a prerequisite of developing the reading literacy is enjoying reading and having internal need to read.

• Literal understanding - reading literacy is based on the skill to decode written texts and to establish compehension while using the so far acquired knowledge and experience.

• Deduction and evaluation – a person with reading literacy is able to deduce and draw conclusions from a text and to see the texts from different viewpoints.

• Metacognition – a selfregulation, a skill to reflect on one's reading and to set the aim of the reading, is a part of reading literacy. An individual is able to choose a proper text and the appropriate reading method according to his target. An individual is able to understand the text and to choose a strategy to improve the understanding and to overcome difficulties with regards to the content and complicated formulations.

• Sharing – a person with reading literacy is ready to share his experiences and understanding with other readers

• Application – a person with reading literacy uses reading for his selfdevelopment and he benefits from it in his further life

At present, the reading literacy is considered to be poor and inefficient, which results in problems with studying, education decline and consequently also in problems when looking for a job. The reading literacy, as a prerequisite of continuous education and successful integration of an individual into society, has become the core of the curriculum in all developed countries. The reading literacy has become a subject of many international comparative researches (e. g. PIRLS, PISA, etc.). Teaching

reading traditionally belongs to priorities of the Czech educational system (KRAMPLOVÁ, I. (2012) Národní zpráva PIRLS 2011. Praha: Česká školní inspekce).

This research focuses on the chosen topical issues of the reading literacy and their impact upon the educational process from the primary school teachers' viewpoint. It has been a one-year grant project. Its empirical investigation started in March 2013

and will last till the beginning of the year 2014. The empirical investigation focuses on primary school teachers. It has been done by means of an anonymous questionnaire and observations in the lessons. The investigation has been carried out in primary schools chosen by simple random sampling method within the territory of the Czech Republic. At present, 500 primary school teachers represent the research sample. Together with the research a short questionnaire for pupils has been created to get a complex view of the issue. We aim at acquiring both the teachers' and the pupils' viewpoint.

The pedagogical research of the reading literacy is based upon quantitative as well as qualitative research methods.

The main research tool has been the non-standardized questionnaire method, which enables us to collect the required information from a large number of respondents within a relatively short time. The questionnaire has been tested by means of preresearch the aim of which was to eliminate unsuitably formulated questions or questions that require unknown information.

The emphasis has been put on stylistic accuracy and grammatical correctness of the questions. Since the questionnaire is non-standardized, closed questions with a choice of an answer, then the questions with a choice of several answers, and also open questions enabling free answers have been included. The questionnaire can be divided into 2 parts – the informative – providing information about the teachers' education, gender, number of years they have worked as teachers and their satisfaction with their job. The second part of the questionnaire represents the reading literacy investigation. As regards the reading literacy questionnaire content, it dealt with reading literacy development means/methods/tools, the way they are used, genre preferences and school and class libraries availability. The questionnaire consisted of 40 items. When creating the questionnaire we intended to make up a questionnaire that would be comprehensible, with clear questions whose number must be adequate. So far 431 teachers (412 women, 19 men) from all parts of the Czech Republic, including small class teachers, have participated on the research. The respondents are teachers with different length of teaching experience. The research is still in progress (GAVORA, P. (2010) Úvod do pedagogického výzkumu. Paido, Brno. CHRÁSKA, M. (2010) Metody pedagogického výzkumu. Grada, Praha).

The aim of the investigation focused on teachers has been to find out to what extent they are informed in the researched area, how they view the importance of the reading literacy in the context of primary education, what teaching strategies they apply in order to develop the reading literacy not only in the literature education lessons, but also in other school subjects.

One of the major qualitative methods has been observation, which means a purposeful, systematic and methodical perception of phenomena and processes which aims at discovering essential connections and relations in the field. The observation has been properly organized, has followed the planned schedule and the observed phenomena and processes are properly registered, which will serve as a complementary method of the questionnaire investigation.

The final stage of the observation method is the analysis of observation results. The aim of this method is not only the description and registration of facts, but also speci-

fying or formulating hypotheses. Within the presented research we will apply direct and short-term observation (GAVORA, P. (2010) Úvod do pedagogického výzkumu. Paido, Brno. CHRÁSKA, M. (2010) Metody pedagogického výzkumu. Grada, Praha).

4 Empirical Analysis

The collected data have been ordered and frequency tables have been created to get their systematic overview, on the grounds of which the graphical charts have been created.

The basic data sorting have been done due to so called "dashed line method," where all the collected data have been registered and consequently their occurrence in the whole research has been recorded. Their relative frequency, showing how often their occurance in certain cathegory is, has been expressed. The acquired data have been put into a graph.

For this purpose frequency histograms have been used. In the charts, x-axis represents the measuring value and y-axis represents its frequency. To describe the structure of the chosen item, pie chart has been used.

To work out the position characteristics, the arithmetic average has been used, mainly because of its simplicity, which makes the data for readers easy to understand. It also enabled us to deduce further important relationships. Last but not least we have created variation widths, to assess the value dispertion and to set the difference between the biggest and the smallest measured value.

Our contribution summarizes the most important findings to provide a complex view of the reading literacy issue (GAVORA, P. (2010) Úvod do pedagogického výzkumu. Paido, Brno. CHRÁSKA, M. (2010) Metody pedagogického výzkumu. Grada, Praha). The acquired results can be divided into 4 basic areas

- school and class libraries and their usage.
- teaching materials
- types of texts used in the lessons
- reading literacy developing activities

First, we will focus on the way school and class libraries are used in the education. Our research has revealed, that for 328 teachers (76%) a school library and for 237 teachers (55%) a class library is available. Class libraries are used more often (see fig. 1 and fig. 2).

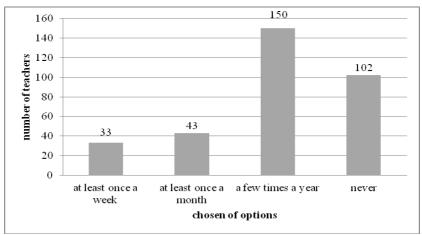
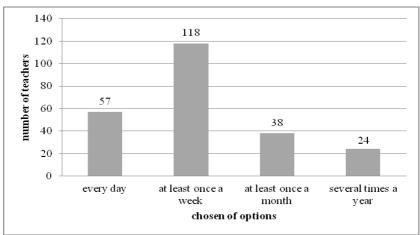
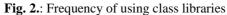


Fig. 1.: Frequency of using school libraries

Only 9% of the teachers have said, that they visit the library at least once a week. 32% of the teachers never use it. Class library is used daily by 24% of the respondents and once a week by 50% of the respondents (see fig. 2).





In the next question we were interested in the structure and number of the materials available in the libraries. It is well-known, that there are considerable differences among Czech schools in this area. The teachers should say how many books their libraries contain. The following table summarizes the collected informations.

	200 and less	201 - 500	501 - 1000	1000 - 2000	2000 and more
City	4 %	15 %	36 %	32 %	13 %
Village	17 %	19 %	31 %	30 %	3 %

Table. 1.: Differences between cities and villages

We have found out that the most common sources in the Czech school libraries are children books, magazines for children, textbooks, reading books. Dictionaries, special literature for teachers, non-fiction, CDs and DVDs do not occur that often.

Further in the research, the means teachers use in the education, have been analyzed. As the research has shown, the most frequently used means in reading and literature education in primary schools are reading books, children books (see fig. 4). Less frequented are on contrary work sheets and the least used means is PC. Among the items, in the chart indicated as "others," the teachers named interactive boards, school magazines and some other texts.

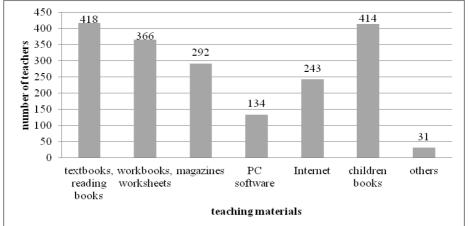


Fig. 3.: Teaching materials used

Another part of the research focused on the types of texts used in the lessons.

Types of texts	Number of teachers	Number of teachers %
fables and fairy tales	405	94
Fiction	422	98
longer books with more chapters	353	82
Poems	427	99
theatre plays	310	72
non-fictions	164	38
instructions and handbooks	56	13
Graphs	112	26

As the table above shows the most popular texts for developing reading literacy of pupils are poems, fiction, especially fairy tales and fables. On contrary, instructions, handbooks and graphs are less popular.

Using non-fiction is very rare.

The last issue analyzed in our research was activities used in the lessons to develop reading literacy of the pupils. The teachers were asked to choose the way they help pupils with texts.

Activities	Number of teachers	Number of teachers %
teachers read aloud	414	96
pupils are asked to read aloud	431	100
pupils read quietly	422	98
pupils are given time to choose books themselves	280	65
teachers teach or perform read- ing at a glance and information decoding strategies	181	42

Table 3. : Reading activities

As the table shows, reading aloud has been identified as the most frequented reading literacy developing activity. Considerably less respondents, only 181 of them, teach or perform reading at a glance and information decoding strategies.

In the next part of the research, the teachers have been inquired about activities required from the pupils to be done in their literature lessons.

Activities	Number of teach- ers	Number of teachers%	
Information looking up	418	97	
Identifying the main thought	314	73	
Explain their understanding of the text	384	89	
Comparing with their own experience	267	62	
Comparing with other texts	280	65	
Predicting the conclusion of the text	220	51	
Conclusion drawing and general- ization	371	86	
Describing text style and structure	155	36	

The teachers mostly want their students to look up information – this answer has been chosen by 97% of the teachers. Pupils are also required to explain their understanding of the text and to draw conclusions (89%) or to make generalizations (86%). Only 36% of the teachers want their pupils to describe the text style and structure. In

comparison to the international comparative research PIRLS 2011 Czech teachers do not use all existing means. In the new EU countries 4 methods and in the old EU countries 3 methods are used. About 30% of teachers in the old EU countries use 7 methods and the same number of teachers in the new EU countries uses 6 methods (KRAMPLOVÁ, I., POTUŽNÍKOVÁ, E. (2005) Jak (se) učí číst. Praha: TAURIS).

The second part of the whole research – observation – has been done in literature lessons. We focused on textbooks that have been chosen, teaching strategies and reading literacy developing activities used. We were also interested in homework that might encourage pupils' interest in reading.

At this point we would like to provide a view to this topic through pupils. In questionnaire investigation which is part of research project we focused on 5th grade pupils. One part of the questionnaire was focused on reading lessons. There were several statements and children were asked to fill how much they agree with these statements about their reading lessons. (see Table 5).

Statements	Agree a lot	Agre e a little	Disagree a little	Disagree a lot
I like what I read about in school	12 %	50 %	30 %	8 %
My teacher gives me interesting things to read	10 %	32 %	45 %	13 %
My teacher gives me interesting things to do	17 %	26 %	36 %	21 %

Table 5.: Statements about reading lessons

Table 6 presents results of activities in school which can improve reading literacy. Our research shows that only 30 % children read silently for their own every day (or almost every day), 38 % children never (or almost never) read books that they choose themselves.

Table 6.: Activities in school

Activities	Every day or almost every day	Once or twice a week	Few times a month	Never or almost never
I read silently on my own	30 %	28 %	18 %	24 %
I read a books that I choose myself	16 %	22 %	24 %	38 %

Following all findings acquired within our direct, short-term observation, we have come to a conclusion, that Czech teachers keep using traditional and approved strategies. When speaking about books read in literature lesson, The Six Bullerby Children or Honzíkova cesta are the most read books. Our observation findings confirmed the data collected in the questionnaire. Reading aloud, followed by teacher's explanation of the unknown words remain to be the most popular strategies. In some cases, quiet reading for an exactly given time has been introduced. When practising reading literacy skills, pupils were asked to look up information in the text, formulate the core of the content or explain the read information. In contrast to the questionnare findings, application of own experience of the pupils was more often recquired. Homework is assigned and checked every day. Unfortunately, teachers usually insist on reading classical literature without any effort to encourage pupils' interest in reading by activating creativity of the pupils.

5 Conclusions

Our contribution has presented results of our research into chosen topical issues of reading literacy and their impact upon pedagogical educational process from primary school teachers' viewpoint. Its target was to analyze the current situation with reading literacy of young pupils. The central point of the project represented is its empirical investigation done on the grounds of an non-standardized questionnaire. The teachers questionnaire enabled us to observe strategies used in reading and literature education. Besides, we were given an opportunity to judge school and class libraries and of course to deal with reading literacy developing activities used in Czech schools. Our research findings have revealed that despite a wide range of methods and means available, Czech teachers prefer traditional concept of teaching.

The research managed to create a profile of the current pupil-reader from the primary teachers' viewpoint.

Referring to the data from the pupils questionnaire we can say that reading is not a very popular leisure time activity with primary school pupils. Most of the respondents do not even have a bookshelf at home. Current readers seem to prefer adventurous and fantasy literature. Printed literature dominates over the electronic one.

The research has provided an overview of the current situation with regard to teaching reading and literature education in primary schools. We have also managed to create characteristics of subjective as well as objective factors of reading literacy. The research has formulated conditions for efficient development of the primary school pupils'reading literacy level and it can inspire teachers to evaluate the school education programmes with respect to the development of pupils'reading literacy.

The text provides up-to-date conclusions and suggestions for the optimisation of teaching literature education, emphasizing the application of activisation methods and forms of work that lead to the reading literacy development of primary school pupils.

Coming from the teachers questionnaires, primary school teachers consider reading literacy development to be one of the educational priorities. However, they have not started using all existing means for increasing the efficiency of their work in this area.

This contribution has been created due to the student grant project PdF_2013_033 Reading Literacy in the Educational Process from Primary School Teachers' View-point and the project PdF_2013_023 Literature and Media in the context of the development in reading literacy amongst elementary school children.

References

1.BARRY A. L. (2002) Reading strategies teachers say they use. Journal of
adolescent and adult literacy. Available from:
http://moormangb.ced.appstate.edu/5534_f07/unit2/barry_strategies.pdf

2. EACEA P9 EURYDICE. (2011) Teaching Reading in Europe: Contexts, Policies and Practices. Brussels: Education, Audiovisual and Culture Executive Agency).

3. GAVORA, P. (2008) Ako rozvíjať porozumenie textu u žiaka. Enigma, Praha.

4. GAVORA, P. (2010) Úvod do pedagogického výzkumu. Paido, Brno.

5. GEJGUŠOVÁ, I. (2012) Vliv mimočítankové četby na rozvíjení čtenářských aktivit žáků. Metody a formy práce ve výuce mateřského jazyka. Hanex, Olomouc.

6. HYPLOVÁ, J. (2010) Využití projektového vyučování k rozvoji čtenářské gramotnosti žáků základní školy. Pedagogická fakulta Ostravské univerzity v Ostravě, Ostrava.

7. CHRÁSKA, M. (2010) Metody pedagogického výzkumu. Grada, Praha.

8. KRAMPLOVÁ, I., POTUŽNÍKOVÁ, E. (2005) Jak (se) učí číst. Tauris, Praha.

9. KRASHEN, S. D.(2004) The power of reading. Heinemann, USA.

10. KUCER, S. B. (2005) Dimension of Literacy. A Conceptual Base for Teaching Reading and Writing in

School Settings. London : LEA, London

11. METELKOVÁ – SVOBODOVÁ, R., HYPLOVÁ, J. (2011) Strategie rozvíjení čtenářské gramotnosti v čítankách 1.stupně ZŠ. Ostravská univerzita v Ostravě, Ostrava.

12. NAJVAROVÁ, V.; NAJVAR, P. (2007) Reading Strategies and the development of Reading Skills of primary school pupils in the Czech Republic. Berlín: DGLS. Available from: http://www.literacyeurope.org/idec-thecommittee/?option=com_content&task=view&id=97&Itemid=47

13. RABUŠICOVÁ, Milada. (2002) Gramotnost: staré téma v novém pohledu. Masarykova univerzita & Nakladatelství Georgetown, Brno.

14. TEMPLE, CH.; OGLE, D.; CRAWFORD, A. FREPPON, P. (2005) All Children Read. Teaching for Literacy in Today's Diverse Classrooms. Pearson Education, Boston.

15. Výzkumný ústav pedagogický v Praze. (2011). Studie k problematice čtenářské gramotnosti v základním vzdělávání – 1. část. [online] Available from: http://clanky.rvp.cz/clanek/c/Z/13691/studie-k-problematice-ctenarskegramotnosti-v-zakladnim-vzdelavani-1 - cast html/

gramotnosti-v-zakladnim-vzdelavani-1.-cast.html/

16. Výzkumný ústav pedagogický v Praze (2011). Studie k problematice čtenářské gramotnosti v základním vzdělávání – 2. část. [online] Available from: http://clanky.rvp.cz/clanek/c/Z/13723/studie-k-problematice-ctenarskegramotnosti-v-zakladnim-vzdelavani-2.-cast.html/

17. Výzkumný ústav pedagogický v Praze (2011). Studie k problematice čtenářské gramotnosti v základním vzdělávání – 3. část. [online] Available from: http://clanky.rvp.cz/clanek/c/Z/13725/studie-k-problematice-ctenarskegramotnosti-v-zakladnim-vzdelavani-3.-cast.html/#a2

Developing a Workplace Cyberbullying Measure

Sam Farley¹, Carolyn Axtell¹, Christine Sprigg¹, and Iain Coyne²

¹ Institute of Work Psychology, University of Sheffield, sjfarley1@sheffield.ac.uk ² Institute of Work, Health and Organisations, University of Nottingham

Abstract. Workplace cyberbullying refers to bullying behaviours repeatedly enacted by employees through communication technologies. Academic understanding of the phenomenon is limited, further research is therefore needed to better understand how it affects employees and how it can be prevented. Such research is currently being hindered by a lack of adequate psychometric measures. The proposed study seeks to address this issue by using Hinkin's (1995) six step measure develop process to create a valid and reliable scale that can assess workplace cyberbullying. This will benefit researchers by providing an appropriate investigation tool with which to conduct future research.

Keywords: workplace cyberbullying; measure development; communication technology

1 Introduction

Cyberbullying is defined as "An aggressive, intentional act carried out by a group or individual, using electronic forms of contact, repeatedly and over time against a victim who cannot easily defend him or herself" (Smith et al, 2008, p. 376). Academic exploration of cyberbullying began during the early 21st century, at a time when high speed internet services were being developed to allow broader access. Like most media coverage of the phenomenon, the predominant research focus has been on cyberbullying perpetrated by children and adolescents. However given the widespread use of communication technologies in organisations, workplace cyberbullying has become a problem for employees and Human Resources departments.

Literature Review

Studies conducted on cyberbullying among children and youths can inform organisational scholars about its nature and how it differs from traditional bullying. There are several unique features of cyberbullying. It can be experienced at home, which has traditionally been a place where victims can escape bullying. Some acts of cyberbullying can also be seen by a much greater audience than offline bullying and the perpetrator is often unaware of the victim's reaction to cyber events (Shariff, 2005). This can lead to a reduction in empathy and misunderstandings over what constitutes bullying behaviour (Bhat, 2008). Research involving child and adolescent samples has revealed that cyberbullying is associated with a range of negative outcomes. It has been related to behavioural effects, such as increased absenteeism (Katzer, Fetchenhauer, & Belschak, 2009) and lower academic achievement (Beran & Li, 2007). Psychological effects have also been identified, including depression (Didden et al, 2009), social anxiety (Juvoven & Gross, 2008) and lower self-esteem (Katzer et al, 2009).

At present, there is a lack of published studies on workplace cyberbullying. Yet the research which has been conducted indicates that it is a problem. Privitera and Campbell (2009) conducted a study on cyberbullying in the Australian manufacturing sector and reported a prevalence rate of 10.7%. Whilst Baruch (2005) found that bullying via email was associated with an intention to leave the organisation, lower job satisfaction and anxiety. Further research is needed as virtual work is becoming increasingly common. Email is the most widely used organisational communication tool and video conferencing has become a popular, cost efficient way of connecting employees. Furthermore, workplace bullying of any kind threatens employee well-being and exerts a financial cost on organisations. A study conducted in 2008 estimated that workplace bullying costs UK organisations £13.75 billion per year (Giga, Hoel & Lewis, 2008). One of the chief reasons for the dearth of research on workplace cyberbullying is the lack of psychometrically sound measurement tools. It has therefore been suggested that future research should concentrate on the development of reliable and valid cyberbullying measures (Menesini & Nocentini, 2009).

The development of a workplace cyberbullying measure will contribute to research and practice. From a research perspective it is important to obtain precise estimates of the prevalence of cyberbullying as over or underestimation can produce Type I or Type II errors (Nielsen, Matthiesen & Einarsen, 2010). Validity is also important as data collected using bullying measures contributes to the design and implementation of intervention strategies (Schat, Frone & Kelloway, 2006). Furthermore, Nielsen et al (2010) have suggested that governments and organisations depend on correct estimates of workplace bullying so that they can budget time and resources to address the problem. The same argument applies for cyberbullying as organisations will need to be aware of whether they have a cyberbullying problem and what needs to be done to address and prevent it.

Proposed Methodology

Prior to data collection a review of the workplace bullying and cyberbullying literature will be undertaken in order to produce a definition of workplace cyberbullying. The provision of a definition is deemed necessary because some individuals lack an understanding of cyberbullying and without a definitional outline, findings would be inconsistent (Tokunaga, 2010). Once a definition has been produced, the six step measure development process outlined by Hinkin (1998) will be followed. Item generation is the first stage in this process and potentially the most important aspect of developing a reliable measure (Hinken, 1995). As cyberbullying is a novel research area, behavioural descriptions will be obtained from populations of interest as recommended by Dawis (1987). These descriptions will be gathered from individuals working in different industries to ensure that the behaviours are relevant to a variety of occupations and organisations. Respondents will be asked to describe examples of cyberbullying behaviour that are consistent with the proposed definition. Items will also be independently generated by searching through journal articles published on cyberbullying.

Following the item generation process, the items will be reviewed by several expert judges to ensure content validity. The judges will assess items for clarity, consistency with the definition and the extent to which they are relevant across jobs and industries, thereby reducing the number of usable items. The usable items will form the basis of the measure, which will be distributed to a relevant sample of employees. Hinkin (1998) recommends a sample of over 200 for factor analysis purposes. Exploratory factor analysis (EFA) and reliability analysis will then be conducted to further reduce the number of items.

If the initial sample exceeds 400 the responses can be randomly split into two and confirmatory factor analysis (CFA) can be conducted on the other half of the sample. If not, a further sample of 200 will be obtained for CFA. To establish convergent validity, the questionnaire will be completed alongside a current cyberbullying measure, such as the cyber Negative Acts Questionnaire. The final step in the process is to repeat the administration on a new sample. This is especially pertinent if items have been deleted during EFA. The new data will then be subjected to CFA, reliability analysis and checks for divergent and convergent validity. These analyses will provide confirmation of whether the measure possesses sufficient reliability and validity to be used in future research.

Conclusions

The implementation of communication technologies in modern organisations has seen the rise of a new form of workplace bullying, termed workplace cyberbullying. At present little research exists on the phenomenon, academic exploration is therefore needed to understand the impact it has on employees and organisations. One of the barriers to such research is the absence of a valid and reliable measure. This paper sets out how a psychometrically sound workplace cyberbullying instrument will be developed for the benefit of academics and practitioners.

References

Baruch, Y. (2005). Bullying on the Net: Adverse behaviour on email and its impact. *Information and Management*, 42, 361-371.

Beran, T., & Li, Q. (2007). The relationship between cyberbullying and school bullying. *Journal of Student Wellbeing*, 1(2), 15-33.

Bhat, C. S. (2008). Cyber Bullying: Overview and Strategies for School Counsellors, Guidance Officers and All School Personnel. *Australian Journal of Guidance and Counselling*, 18, 53-66.

Dawis, R. V. (1987). Scale Construction. *Journal of Counseling Psychology*, 34(4), 481-489.

Didden, R., Scholte, R. J., Korzilius, H., De Moor, J. H., Vermeulen, A., O'Reilly, M., et al. (2009). Cyberbullying among students with intellectual and developmental disability in special education settings. *Developmental Neurorehabilitation*, 12, 146–151.

Giga, S., Hoel, H., & Lewis, D. (2008). *The Costs of Workplace Bullying*. Research Commissioned by the Dignity at Work Partnership: A Partnership Project Funded Jointly by Unite the Union and the Department for Business, Enterprise and Regulatory Reform.

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management*, 21(5), 967-988.

Hinkin, T. R. (1998). A Brief Tutorial on the Development of Measures for Use in Survey Questionnaires. *Organizational Research Methods*, 1, 104-121.

Juvonen, J., & Gross, E. F. (2008). Extending the school grounds?-Bullying experiences in cyberspace. *Journal of School Health*, 78, 496- 505.

Katzer, C., Fetchenhauer, D., & Belschak, F. (2009). Cyberbullying: Who are the victims? A comparison of victimization in Internet chatrooms and victimization in school. *Journal of Media Psychology*, 21, 25–36.

Menesini, E., & Nocentini, A. (2009). Cyberbullying definition and measurement: Some critical considerations. *Zeitschrift für Psychologie/Journal of Psychology*, 217, 230–232.

Nielsen, M. B., Matthiesen, S. B., & Einarsen, S. (2010). The impact of methodological moderators on prevalence rates of workplace bullying. A meta-analysis. *Journal of Occupational and Organizational Psychology*, 83, 955-979. Privitera, C., & Campbell, M. A. (2009). Cyberbullying: The new face of workplace bullying? *CyberPsychology & Behavior*, 12, 395-400.

Schat, A. C., Frone, M. R., & Kelloway, E. K. (2006). Prevalence of workplace aggression in the U.S. workforce: Findings from a national study. In E. K. Kelloway, J. Barling, & J. J. Hurrell (Eds.), *Handbook of workplace violence*. Thousand Oaks, CA, 47-89.

Shariff, S. (2005). Cyber-dilemmas in the new millennium: School obligations to provide student safety in a virtual school environment. *McGill Journal of Education*, 40, 467-487.

Smith, P., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., & Tippett, N. (2008). Cyberbullying: its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry*, 49, 376-385.

Tokunaga, R. S. (2010). Following you home from school: A critical review and synthesis of research on cyberbullying victimization. *Computers in Human Behavior*, 26, 277-287.

Exploring the *Gesamtkunstwerk* concept in the design of the theatrical spatial experience. A constructivist

Thaleia Grigoriadou,

approach.

PhD candidate - Lecturer in yearly contract, School of Architecture A.U.Th., Greece, (PhD supervisor: Dr. Fani Vavili, A.U.Th.)

thalia_gr@post.harvard.edu

Abstract. According to the Wagnerian tradition, Gesamtkunstwerk refers to the creative fusion of multiple types of expertise from the worlds of art, theory, philosophy as well as technology. From Romanticism, to modern and postmodern times, this concept seems to occasionally reflect society's goals and artistic aspirations; yet, in modern times it has frequently been associated with totalitarianism advancing ideals of national identity through total clarity of principle. In this historical retrospective, theatrical and performance space is a field where Gesamtkunstwerk has been distinctively as well as earnestly implemented. Today, this integration is further traced in the creator's use of the computer as a meta-medium, synthesising all media within a single interface, sharing time and space with the performers' physical presence. Is there an actual need for a unifying synthesis when postmodern beliefs propound differentiation and fragmentation, often reaching a design extravagance? Does artistic plethora provide enough space for various artistic minds to participate towards the common performance work? Does it make for an artistic democracy or does it shrink creation?

By employing *Gesamtkunstwerk* as the main theoretical frame of this work and Constructivism as the key methodological paradigm, I will explore its contemporary expressions in the interplay between artistic practices for collaborative performance, as well as suggest that it informs current performance design in challenging ways. The expected contribution of this research is the emergence of some new ways of understanding the history and contemporary importance of *Gesamtkunstwerk* and how it affects a designer's work among co-creators in the eventual making of the theatrical spatial experience.

Keywords: Gesamtkunstwerk, total work of art, constructivism, meta-medium, synthesis/union of arts

1. Introduction

1.1 General performance-related definitions

Initially, one should attempt to formulate a clarification of the principal terms that are used in this research such as 'performance', 'performers', 'spectacle', so as to facilitate the understanding of the further analysis. It will not take long to realise that the above task is greatly difficult since these terms are of a less concrete meaning. To begin with, *performance* can be just about anything that is entertaining. People engaged in its practice can be called actors, dancers, street performers/musicians, buskers, minstrels and they may do acrobatics, clowning, comedy, fire eating, fortune-telling, juggling, magic, mime, musical performance, puppeteering and several other artistic activities. The boundaries between entertainment and art, between audience and performers, as well as between actual performance and wider social events - even though today performance studies is not an emerging discipline but a concrete academic field - still seem obscure.

1.2 Gesamtkunstwerk

A general translation of the term *Gesamtkunstwerk* is given as the total work of art [1], ideal work of art [2], universal artwork [3], synthesis of the arts, the allembracing art form or total artwork. A work of art that makes use of all or many art forms or strives to do so. A performance-related explanation of the term, that suggests its collaborative qualities, describes it as "the desire for and practice of combining various art forms into a whole, such as performances that combine text, visual arts, music, dance, and architecture" [4]. Today, it's a notion that mostly serves as a generally admissible word to characterize widely disparate efforts to unite the arts in spirit and in practice.

1.3 Gesamtkunstwerk in Architecture

While it is difficult to discover when the notion of the *Gesamtkunstwerk* was first employed from the point of view of a building and its contents, it is not unidentified in the actual architectural context. It originally signified a holistic meaning, the situation where the architect is responsible for the design and/or overseeing of the building's wholeness: shell, accessories, furnishings, and landscape [5]. Already during the Renaissance, artists such as Michelangelo saw no strict division in their tasks between architecture, interior design, sculpture, painting and even engineering, supporting the union of arts approach in one man's/designer's direction. Distinctive modern paradigms of such architects are Frank Lloyd Wright (1867-1959) and Alvar Aalto (1898-1976).

2. Literature Review

2.1 Various expert theories on the *Gesamtkunstwerk* idea 19th century. Wagner, the later Impressionists and the inner meaning

Although general conceptions of unity, integrity and completeness of the artwork are found in antiquity, the emergence of the aesthetic as a theoretical category is found in late eighteenth-century German thought and officially introduced by German composer and theatre director Richard Wagner in the nineteenth century. According to Wagner's beliefs, the abstract form of the drama should dominate the subordinate elements of music, words and other performing arts that abandon their exclusive identity to find a higher place in their synthesis for the new society [6]. Wagner raises the issues regarding approximating beauty, nobility and freedom that can only be attained with social revolution, when further speaking of a public artwork and of a "public fully free and independent when face to face with Art" [7].

In performance space, the idea of the inner soul of the whole, originally expressed in music by Wagner, was to be later explored in the visual realm by the Impressionists and Adolphe Appia. They too perceived the complete composite art work as one requiring a sole designer/director, a *one* creative mind, to bring all the elements of the stage into harmony. Like the Romantics, the Impressionists found the deeper meaning of performance by removing the surface details and reaching more abstract patterns of mood and atmosphere. They've heightened the visual elements of the play by developing a flexible setting and lighting system that would make the moving actor "expressive to the eye" [8].

Painter and theorist Wassily Kandinsky has worked on abstract performance compositions. *Der gelbe Klang (The Yellow Sound)* is one scenic experiment of 1909, along with other creators, where he practically expresses theories of colour 'synesthesia' and 'independence' in the union of the arts: all the artistic elements (including colour and decoration) would be equally important... Kandinsky hoped to achieve a new kind of inner unity, since all the media involved would be used according to the artist's "inner necessity" of expression [9].

20th century. Bauhaus: the modern approach to *Gesamtkunstwerk* and *Totaltheater*

In the modern époque the arts were brought closer together in many ways. The major art movements have contributed new ideas and approaches for crossing the traditional lines of the separate disciplines, seeking for the synthesis of artforms. Impressionism in painting, music and poetry, Cubism in painting and poetry, Futurism in poetry, graphics, music, film, painting, and total performance, or Expressionism combined nearly every artform [10].

A particularly modern idea of *Gesamtkunstwerk* -one that approached art and architecture much closer together- emerged with the Bauhaus school that assembled design, art and craftsmanship in common curriculum. It was first established in Weimar in 1919 by Walter Gropius who expressed that artists and architects should also be craftsmen and that they should be aware and use different materials and artistic mediums, including industrial design, theatre/costume design, music. Design was not necessarily being the work of a single hand and collaboration of expertises was strongly supported [11].

The critical blend in Gesamtkunstwerk. Fragmentation and collage

German poet and theatre practitioner Berthold Brecht supports the emergence of an openly divided and critical blend of art in theatre space that could be called *Gesamtkunstwerk*: he aims to develop a model of artistic unity through the art of editing, while he tries to re-imagine the initial concept without the Wagnerian organic aesthetics. "So long as the expression *Gesamtkunstwerk* means that the integration is a muddle, so long as the arts are supposed to be 'fused' together, the various elements will all be equally degraded, and each will act as a mere 'cue' to the rest. The process of fusion extends to the spectator who gets thrown into the melting pot too and becomes a passive (suffering) part of the total work of art. Witchcraft of this sort must of course be fought against..."[12]

A more contemporary expression of the Brechtian idea of independence and fragmentation is the one of *collage*, a practice that is associated most often with the visual arts, nonetheless, is noticeable in the time-based and performance space as well: collage is a principle organizing strategy in the work of Elizabeth LaCompte and the Wooster Group, the theatre pieces of Robert Wilson, the choreography of Pina Bausch and Merce Cunningham [13], as well as in the contemporary Greek paradigms of Dimitris Papaioannou and Constantine Rigos performances. '*Nowhere*' ($\Pi ou \partial \varepsilon v \dot{\alpha}$) is a contemporary work directed in 2009 by D. Papaioannou, being presented at the

National Theatre of Greece, utilizing the new -at the time- stripped mechanisms of the theatre's main stage. It could be seen as a spatial experience-specific project, both for creators and audience, where performers 'count and weigh to the dimensions and capabilities of the space', testing human as well as artificial material, all in the framework of a specifically designed performance which cannot occur anywhere else: "*Nowhere* is a project for the space of the theater scene. (A project for) the spacemachine that continuously changes and is determined by the human presence to signify countless locations, while it is designed to be a non-place." [14]



Fig. 2. '*Nowhere*' (2009), National Theatre of Greece. Directed by D. Papaioannou, Stage design associate: Z. Xagoraris. (Source: http://www.n-t.gr/el/events/poithena/)

Contemporary times: evolution of *Gesamtkunstwerk* as "the medium of phantasmagoria"? *Meta*-issues and critique on the contemporary spectacle and society

Contemporary scenarios of *Gesamtkunstwerk*'s union have often provoked the characterizations of the *techno-utopia* notion: issues of infinite -pointless? – regathering of the dispersed forms of individual arts emerge and the artwork concept appears in debates on the autonomy of media and on the relation of art to life, "displaying an aesthetic ambition to borderlessness that encompasses a variety of artistic, political, and (even) metaphysical levels... Cyberspace continues to blur the distinctions between art and commercial spectacle" [4]. The *Gesamtkunstwerk* idea of the total union seems to revive in today's multi-media events, technology shows and experiments, and generously provides an extended (infinite?) space for new artforms to meet and co-create.

On the other side, one of the most trenchant critiques on the Wagnerian concept, further implemented for its contemporary presence, comes from German sociologist, musicologist and critical theorist Theodor Adorno. Adorno describes Gesamtkunstwerk as a rationally constructed artificial entity that ignores the social conditions necessary for the survival of the whole. Moreover, he talks of inner requirements of the artistic material that are not met and points out a predetermined plan that deprives the individual creator from free artistic will. Furthermore, he argues that the seemingly unified totality of Wagner -a mere illusion- aims to the eradication of the individual and links Gesamtkunstwerk directly to the totalitarian ideology (Wagner's link to the 3rd Reich fascist ideology) and the regression of spectatorship. Spectatorship becomes a kind of a 'dozing' situation, subject to the creator's psychological control. In this way, 'the process of synthesis is the agent of ideology' [15]. While Gesamtkunstwerk marks the victory of technology and rationality Adorno aligns with Nietzsche in terms of ignoring the strict boundaries placed on disciplines, resisting systematization and condemning the systematic wholeness [16].

3. Proposed Methodology

At this point I would like to clarify the hermeneutical methodology pursued, and refer to the Constructivist paradigm I follow in order to investigate the *Gesamtkunstwerk* case. I intend to explore the multiple socially, experientially based mental constructions coming both from experts and possibly further discover current subjective theories that relate to the *Gesamtkunstwerk* idea in today's life and possibly reveal important information on the design field of the theatrical experience [17].

4. Rationale and early conclusions

Today's creative practitioners seem to have the ambition for performance design experimentation in spaces they choose to engage the spectator. They wish their works to cross the formal boundaries of the disciplines in designing for the spatial experience, with much help from technology innovation being offered. However, fruitful and 'democratic' cooperativeness -in both arts level as well as 'users' level-seems a particularly complicated task. One understands that most of the above raised theories approached *Gesamtkunstwerk* as an abstract or a Romantic ideal concept that, nonetheless, has also reached to serve less or more occult practical social purposes, thus becoming the *topos* of severe critique. It would be 'safe' to admit that one should give some thought before attempting to formulate a contemporary dogma of total artwork collaboration today.

Starting from the literature review formulations and the above observations, I find significant interest in the development of research in two axes and their corresponding dualities:

4.1 Research regarding the multi-experience of the spectator:

Investigation of the duality of hyper-technological, 'overdesigned' culture as opposed to a simpler, more physical spectacle at the place of its birth and presentation. What are the influences and effects in the spectator's spatial experience in each case? I further address contemporary issues of technology 'competitiveness' to the natural human presence (the wider philosophical and ideological battleground of "humans versus machines"), in response to the paradigm of applications of digital and cinema projection in the field of performance space. Among others, I'll be referring to M. W. Smith's insight for a 'meta-organic performance' and the return to the natural [6], as well as to music/performance practitioner John Cage's idea of the audience as a score of individuals instead of a group and his 'anti-dozing' aspirations.

4.2 Research regarding interactivity in the experience of the creators' cooperativeness:

Investigation of the duality of collaborative teamwork abilities and skills as opposed to individual inspiration and personalized creation. The synthetic process for creating performance space, the "roles" and the relationships' dynamics. Reference to the 'one-man-work' paradigm of *encenadores* (directors-designers of the 1970s) and its new generation expression of the *artist-as-subject* idea.

Gesamtkunstwerk originated from the exploration of an inner -social- meaning of the artwork. One might say that the reappearing tendency towards the Gesamtkunstwerk is associated with the disruption of social, political, artistic values of today. In this sense today, "the Gesamtkunstwerk is a social *Reformwerk*" [12]... yet, still searching to serve an inner meaning? This research will try to reveal some new ways of understanding the contemporary importance of Gesamtkunstwerk, its new meaning as designers' and society's "tool", and how this idea might affect a designer's work among co-creators in the eventual making of the theatrical spatial experience. As neoDadaist painter Robert Rauschenberg delivers in the 'Art of Assemblage' [18], each material (creators/designers use) has its content and its independence from the meaning. Maybe the discovery of inner meanings should be left to audiences, even better, to experience.

5. References

- 1. Millington B., Warrack, "Gesamtkunstwerk", http://en.wikipedia.org/wiki/Gesamtkunstwerk
- 2. Oxford English Dictionary, 1972 ed., s.v. "Gesamtkunstwerk"
- 3. Artlex Art Dictionary, "universal artwork", http://www.artlex.com/ArtLex/U.html#anchor1538384
- 4. Finger, A. and Follett, D. (eds.) (2011) *The Aesthetics of the Total Artwork: On Borders and Fragments*, The Johns Hopkins University Press. Review by R. J. Goebel
- 5. Vidalis, M. A. (2010) Gesamtkunstwerk- 'total work of art', *Architectural Review*. http://www.greekarchitects.gr/en/architectural-review/gesamtkunstwerk-id3185
- 6. Smith, M. W. (2007) *The Total Work of Art: From Bayreuth to Cyberspace*, London: Routledge *The Urge for Totality* Review by Alan Ackerman

- Wagner R. (1849) Art and Revolution, W. A. Ellis (trad. in English), *The Art-Work of the Future, Richard Wagner's Prose Works*, Volume 1, Pages 23-65, published in 1895. Original Title Information: *Die Kunst und die Revolution*. Published in 1849, Sämtliche Schriften und Dichtungen: Volume III, Pages 8-41. http://users.belgacom.net/wagnerlibrary/prose/wagartrev.htm
- 8. Kernodle, G. R. (1954) Wagner, Appia, and the Idea of Musical Design, *Educational Theatre Journal*, Vol. 6, No. 3, The Johns Hopkins University Press, p. 226
- Crawford, J. C. (1974) Die glückliche Hand: Schoenberg's Gesamtkunstwerk, Musical Quarterly 60: 583-601; p.590
- 10. Hines, T. J. (1991) *Collaborative form: studies in the relations of the arts*, Kent State University Press, London
- 11. Wittick, A. (1977) Encyclopaedia of Modern Architecture. Thames & Hudson
- Heimbecker, S. (2008) HPSCHD, Gesamtkunstwerk, and Utopia. *American Music 26* (4): 474-498 1967, p. 494
- 13. Copeland, R. (2002), Merce Cunningham and the Aesthetic of Collage, *TDR: The Drama Review* 46: 11–28., p.11
- 14. National Theatre of Greece. http://www.n-t.gr/el/events/poithena
- 15. Adorno, Th. (1954) In search of Wagner, Verso edition (2005), London
- Bauer, K. (2005) Adorno's Wagner: History and the Potential of the Artwork, *Cultural Critique* 60, 68-9, p. 83
- 17. Cuba E.G., Lincoln Y. S. (1994) *Competing Paradigms in Qualitative Research*. In N. K. Denzin &Y. S. Lincoln (Eds.), Handbook of qualitative research (pp.105-117). Thousand Oaks, CA: Sage
- 18. Seitz, W. Ch. (1961) *The Art of Assemblage*. New York: Museum of Modern Art; distributed by Doubleday, Garden City, NY.

Mapping the obstacles to effective governmental and nongovernmental action against drug trafficking: the sociopolitical context in Serbia

Darja M. Koturovic,

¹ 1 University of Sheffield, UK; South East European Research Centre, Greece

dkoturovic@seerc.org

Abstract. Criminal groups nowadays operate as loosely-organized networks of cells, interrelated through diverse transnational networks enabling greater adaptability, wide-scale recruitment and resilience to law enforcement efforts. Globalization, grand corruption and weak governance facilitated empowerment of organized crime groups in Southeast Europe. Multiple impact of organized crime is reflected in the community fear, a lack of trust in the state institutions and generation of vast financial resources commonly used to fund other criminal activities. Aiming to address this phenomenon, Serbian government undertakes significant reforms and a plethora of domestic and international activities to suppress drug trafficking, but the effects of these actions have not yet been assessed. In view of the impact of the social context on the emergence of criminal networks, social network analysis is deployed to map the network of state institutions designed to address drug trafficking groups and identify potential structural holes. The study utilizes elite interviewing in a sample of 50 state officials and civil society representatives to explore the effectiveness of state actions and the influence of a particular social context in this process.

Keywords: drug trafficking, criminal networks, social network analysis, Balkan route, corruption

1 Introduction

Transnational crime networks appear to be major beneficiaries from globalization, the use of IT in international financial exchange and political and economic transition in South East Europe (Trifunovic, 2008; Shelley, 2002; Williams, 2001). Research in different transition countries supports that widespread corruption associated with weak governance and state institutions, facilitates proliferation of organized crime and reinforces illicit activities of transnational criminal organizations (Michaletos, 2009; Shelley, 2002; Vasic, 2010).

The boundaries of conventional international cooperation in the field of law enforcement and judiciary have changed, mainly due to swift mobility, significant diversity of structure and communication methods of criminal organizations. In this line, the fight against organized crime and drug trafficking groups remains one of the most significant priorities of the regional cooperation because of the consequences and high risks for the society associated with this phenomenon. The multiple impact of organized crime is reflected in general fear in the community, a lack of trust in the state institutions, generation of huge financial resources commonly used to fund other criminal activities such as terrorism, and a serious threat for different countries. Thus, the phenomenon of drug trafficking in SEE directly threatens the interests of the region as well as the broader international community.

The Government of Serbia has made the fight against drug trafficking groups and corruption top priorities. These priorities are an attempt to send a clear message, both domestically, that there will be no impunity for such crimes, and internationally, that Serbia will act as a security partner capable of preventing spreading of crime into Europe. The effects of governmental actions have not yet been assessed, but there are indications that an important strike to drug trafficking groups is in progress. In the period of extensive legal reforms performed in Serbia, coupled with strong emphasis on regional collaboration in drug trafficking suppression, further knowledge on functioning of criminal networks and adequate mapping of the affected institutions could offer significant information to future policy making. This research explores certain political and social difficulties in Serbia, which hinder effective work of the relevant institutions and successful state-to-state cooperation in combating drug trafficking.

2. Literature Review

Modern conceptualizations support that traditional structure of organized crime groups such as the top-down hierarchy of mafia, has been replaced by business-like structures that operate as loosely-organized networks of cells (Finckenauer, 2005; Morseli, 2009; Paoli & Fijanoau, 2004; Shelley, 2005). Consistent with these standpoints, social network analysis is often utilized as an approach to study 'dark networks' (Garay Salamanca, Salcedo-Albarán, & de León-Beltrán, 2010; Sparrow, 1991; Williams, 2001; van Duyne, 2003). Research confirms that fluid network structure provides drug trafficking groups with numerous advantages including adaptability, wide-scale recruitment, resilience, as well as capacity for quick learning and innovation (Ayling, 2009; Shelley, 2005). Moreover, networks have greater capacity to exploit new modes of communication and international collaboration than state actors whose activities are based on hierarchical models (Williams, 2001). This type of organization provides greater efficiency, organizational flexibility and hinders law enforcement efforts to identify and position group members (Shelley & Picarelli, 2005). Moreover, transnational criminal groups lack specific ideology, and often use political corruption as a tool for their ends, infiltrating in this way into the political system (Morselli & Giguere, 2006; Shelley, 2005). This is of concern since critical social determinants of organized crime refer to the quality, independence and integrity of the institutions safeguarding the rule of law, including police services and the judiciary.

It has also been put forward that application of social network analysis for the study of the 'dark networks' may be useful to indicate a particular social context that influences the behaviour and 'performance' of actors in such systems (Allie et al., 2008; Garay Salamanca, Salcedo-Albarán, & de León-Beltrán, 2010; Garay & Villaveces, 2009; Klerks, 2001). Following this line of inquiry, the impact of social context was addressed in studies on developing and transitional countries and particularly in Southeast Europe. For instance, the contextual environment in former Yugoslavian states including the ongoing conflicts, embargo and associated state-led criminal ac-

tivities, played an important role in the establishment of strong organized crime groups in the region. In this specific context, multiethnic organized crime groups in the Balkans closely cooperated among each other and other European and transnational criminal groups (Porobic, 2010; Nikolic-Ristanovic, 2004)

Along with the perception of criminal organizations as loose networks, a traditionally apparent presence of legitimate actors in a variety of criminal contexts is of major concern, as it indicates that state actors often closely cooperate with these groups (Antonopoulos, 2008; Block & Chambliss, 1981; Chambliss, 1978; Chevrier, 2004; Di Nicola & Zoffi, 2004; Kleemans & van de Bunt, 2003; McIllwain, 2004; Middleton & Levi, 2004). This implies that corruption is closely associated with growth of crime and should not be perceived as an independent problem. As Shelley argues, compartmentalizing these phenomena limits the capacity to scrutinize and address the diverse forms of transnational crime (Shelley, 2005). Even though there is little empirical research focusing on the link between corruption and organised crime in transition countries, a number of authors emphasize the interdependent nature of these phenomena (Michaletos, 2007; Nikolic-Ristanovic, 2004). For instance, in Serbia, criminal analysts suggest strong linkages between organised crime and public sector corruption, with crime networks greatly relying on corruption to ensure the protection from law enforcement efforts, facilitate money laundering and ensure legal investments based on illicit trade (Nikolic-Ristanovic, 2004; Trifunovic, 2007).

As Warburto (2007) argues, corruption represents a complex, multi factorial phenomenon, closely associated to the fundamental complexity of human social relations. Similarly, Della Porta (1999) emphasizes that corrupt actors must exchange a range of favours in order to protect themselves within the system, referring to this scheme as a market of "corrupt exchanges." Therefore, for corruption to occur, the overall system must be susceptible for such illicit activity. In this line, in view of the aforementioned influence of the social context on the emergence of criminal networks in a given state or a region, Southeast Europe offers a particularly beneficial social context for transnational criminal groups. That is, in transition countries, the large profits generated by drug trade make criminal networks powerful actors who exert influence on politics and business. As a result, they are able to penetrate the highest spheres of government, impact the political agenda and develop close relationships with diverse circles of power, hindering law enforcement investigations. Considering that a number of scientists and security analysts support that "it takes a network to fight a network" (e.g. Walker et al., 2004), the question arises whether and to what extent the state institutions have the capacity to efficiently suppress criminal networks.

Nevertheless, lessons learned show that weakness of the state generates greater vulnerability of the political system to withhold challenges. The association of the growth of organized crime and public sector corruption has been confirmed in a study that used qualitative and quantitative data analysis in a large sample of countries (Buscaglia & van Dijk, 2003). A composite index of organised crime utilized in the study indicated that there is greater prevalence of organised crime in countries characterized by weak rule of law, monopolistic practices and lack of justice efficiency. Moreover, questionable independence of the judiciary and increased political interference in the appointment of civil servants was positively related with higher levels of organised criminal activity. Likewise, a recent study by van Dijk (2007) demonstrated that high levels of corruption and money laundering strongly correlated with weak

rule of law, low justice integrity and lack of responsible institutions.

In the context of affected social bonds coupled with sophisticated interaction methods developed by criminal networks, it appears difficult to destabilize such networks. Finally, the extent of institutionalized corruption and social inertia represent key complicating factors in any attempt to suppress crime in the Balkans. According to Bresson (1999), an institutionalized model of corrupt networks involves "informal social exchanges", within which corruption becomes an extension of "social legal networks" and utilizes a range of personal interests to maintain its function in case of potential socio-political change.

Similarly, the resilience perspective suggests that for combating illicit drug networks, both strategic and operational pathways are necessary (Ayling, 2009). The application of social network analysis therefore appears to be useful for mapping the functioning of the relevant institutions in a given social context. Given the secretive nature of both corruption and organised criminal activities, there is limited ability to produce valid data on the extent of corruption attributable to organised crime. The advantage of utilizing social network analysis lies in its capacity to detect potential structural holes that hinder effective institutional functioning. That is, instead of focusing on the criminal networks as a particular phenomenon, understanding the capacity of the state networks to suppress this activity might be more constructive and offer additional insight regarding the enduring existence of the trafficking groups.

3. Methodology

This study utilizes social network analysis (SNA) to investigate the actions of the government and civil society in addressing drug trafficking networks. This approach results from the common standpoint of contemporary authors that drug trafficking groups function as social networks cellular in structure (Ayling, 2009; Morselli et al., 2007; Shelley, 2002; von Lampe, 2006; Williams, 2001) and that the state should therefore develop a similar network to address this phenomenon.

Social network analysis is often implemented in law enforcement investigations to provide data regarding individual members of illicit drug networks, the relationships among these actors as well as potential network weaknesses. In terms of scientific research on criminal organizations, SNA has been applied to study networks' structure, methods of functioning, interconnectedness of the key players, etc. Analysis of the previous studies using SNA to study criminal organizations indicates significant heterogeneity in the sources of data used, either due to access issues or the specific scope of the research. Data is often retrospective i.e. researchers are able to obtain access only after the state institutions have the necessary information and are willing to provide them. It appears that multiple sources of data could ensure greater validity, that is, a combination of sources could be an effective method to confirm information.

The proposed study is unique in terms of utilization of social network analysis for examination of governmental actions in suppressing drug trafficking networks. Commonly applied for mapping criminal networks, this approach will be deployed in an innovative manner to map the state institutions' network designed to address drug trafficking groups. Through innovative conceptualization of the state institutions as 'actors' or 'nodes' of the network and triangulation of data sources, this study may produce a network map of the governmental bodies aimed at drug trafficking suppression. The mapping of the institutions and positioning their competencies within the network may have the capacity to indicate specific obstacles which can significantly hamper the efforts of the institution. In order to identify such obstacles, betweenness centrality, degree centrality and density measures will be utilized. Patterns of relationship will be explored to detect subgroups, identify central points and the strength of these links. All network nodes and the most frequent links will be analysed in detail. In addition, seemingly loose irrelevant cliques will be explored, since the weak strength of their linkages with other nodes may imply certain obstacles. This process will be utilized to identify structural holes among institutions or obstacles to efficient regional and multiagency cooperation. Concurrently, this method may indicate which institutions are more effective in addressing the problem of drug trafficking, how they are structured and whether such models could be used to inform the system reform. The data obtained through the implementation of social network analysis will be supplemented by the qualitative analysis of the interviews. That is, certain interviews may offer important insight which may further explain or clarify the conclusions. The following research questions will be addressed:

What is the role of the Balkan route as an international grid of illicit drug traffic?

To what extent and in which way the network of state institutions addresses drug trafficking?

What is the strength of the links between the state institutions competent for combating drug trafficking?

What is the role of Serbian CSOs in contributing to systematic documentation and study of this phenomenon?

What are the implications for policy recommendations related to international cooperation in this matter and regional development?

The study will draw upon a variety of official sources, including official court/police statistics; trial reports from Special Court on Organized Crime and Belgrade District Court in drug trafficking and organized crime cases; data on the outcomes of judicial processes in drug trafficking and organized crime cases involving alleged or convicted actors; reports of international organizations; annual CSOs' reports on drug trade, organized crime and money laundering; important journal and newspapers articles by prominent authors in printed and online media. These multiple sources will be analyzed in order to map the existing knowledge on drug trafficking in Serbia and the results of state action in this regard.

Additionally, multiple interviews will be conducted with state officials, the judiciary and civil society organizations. The purpose of the interviews is to obtain more comprehensive knowledge on the work of institutions engaged in drug trafficking suppression, as well as to get more insight into the particular social context in which the institutions function. Elite interviewing will be used to for data collection. The sample will include government officials from the main institutions in the field of combating organized crime and drug trafficking, as well as relevant CSO representatives. Sample size will be approximately 40 to 60 participants. The researcher has already identified potential participants, primarily on the basis of their position. However, it is expected that some participants may indicate other important 'actors' who have not been previously identified (*snowball effect*). This technique will be used as a complementary strategy to obtain more comprehensive data on a specific issue. The study will draw on multiple methods of data collection, data sources and software application. This form of 'triangulation' is deployed to enhance the reliability, interpretability and generalizability of research findings (Robson, 2000).

4. Conclusion

The proposed research is expected to indicate certain political and social difficulties in Serbia, which hinder effective suppression of organized crime and drug trafficking. Even though the Government of Serbia has made the fight against drug trafficking groups and corruption top priorities, the effects of state actions have not yet been assessed. The application of social network analysis may offer important data regarding existing structural holes that have high potential to provide smooth functioning of illicit networks. This may put forward the importance of redesigning the system on the basis of regional and international cooperation and lessons learned. In addition, it may point out alternative ways of transforming law enforcement operations and increase flexibility of organized crime control (Klerks, 2003; Van Dijk, 2007).

References

Aili, E., Malm, A. E., Kinney, J., & Pollard, N. R. (2008). Social Network and Distance Correlates of Criminal Associates Involved in Illicit Drug Production, Security Journal, 21, 77-94.

Antonopoulos, G. A. (2007). Cigarette smugglers: A note on four 'unusual suspects", *Global Crime*, 8 (4), pp.393-398.

Ayling, J. (2009). .Resilience in Criminal Networks, Regulatory Institutions Network, ANU, Centre for Excellence in Policing and Security.

Block, A. and W. J. Chambliss, Organizing Crime (New York: Elsevier, 1981). Bresson, C. J. (1999). Corruption Networks, transaction security and illegal social exchange. In Paul Heywood (Ed) Political Corruption, Blackwell Publishers, Oxford: p49.

Buscaglia, E. & Van Dijk, J. (2003). Controlling organized crime and corruption in the public sector. In: Forum on crime and society, Vol III, no. 1 & 2. United Nations Publication.

Chambliss, W. J. (1978). On the Take: From Petty Crooks to Presidents, Bloomington: Indiana University Press.

Chevrier, E. (2004). The French Government's Will to Fight Organized Crime and Clean Up the Legal Professions, Crime Law and Social Change, (42), 189–200.

Di Nicola, A. & P. Zoffi (2004). Italian Lawyers and Criminal Clients: Risks and Countermeasures, Crime Law and Social Change, (42), 201–225.

Finckenauer, J.O. (2005). Problems of Definition: What is Organized Crime? *Trendin Organized Crime*, *8*(*3*), 63-83.

Garay Salamanca, L.J., Salcedo-Albarán, E., & de León-Beltrán, I. (2010). *Illicit Networks Reconfiguring States*, Social Network Analysis of Colombian and Mexican Cases. Bogotá, Colombia, Metodo Foundation.

Klerks, P. (2001). The network paradigm applied to criminal organizations: Theoretical nitpicking or a relevant doctrine for investigators? *Recent developments in the Netherlands. Connections, 24,* 53–65.

Kleemans, E. and H. G. van de Bunt, The Social Organisation of Human Trafficking, in D.Siegel, H.G. van de Bunt, and D. Zaitch's (eds.), Global Organized Crime: Trends and Developments (Dordrecht: Kluwer, 2003), p. 97–104.

McIllwain, J. S., Organizing Crime in Chinatown: Race and Racketeering in New York City1890–1910 (Jefferson: McFarland and Company, 2004).

Middleton, D. H. and M. Levy, "The Role of Solicitors in Facilitating 'Organized Crime':Situational Crime Opportunities and their Regulation," Crime Law and Social Change,2004 (42), 123–161.

Michaletos, I. (2009). The new international of organized crime in the Balkans. International Analyst Network. www.analyst-network.com

Morselli, C., Giguere, C., & Petit, K., (2007). The efficiency/security trade-off in criminal networks, Social Networks, 29,143-153.

Morselli, C. & Giguere, C. (2006). Legitimate strengths in criminal networks. *Crime,Law & Social Change, 45*, 185–200.

Morselli, C. (2009). Inside criminal networks. New York: Springer.

Nikolić-Ristanović, V.Ž. (2004) Organized crime in Serbia: Media construction and social reaction. u: Meško G., Pagon M., Dobovšek B. (ur.) Dilemmas of Contemporary Criminal Justice, Ljubljana: Fakultet krivično pravnih nauka, str. 685-695

Paoli, L. & Fijnaut, C. (2004). *General Introduction*. In: C. Fijnaut and L. Paoli (eds), 1 -19. Phantoms of Organised Crime, Corruption and Terrorism, Nijmegen: Wolf Legal Publishers.

Porobic, J. (2010). Arms availability and its impact on drug trafficking: A Balkan perspective, Policy Brief Series, no. 13, August, Global Consortium on Security

420

Transformation. Available at: http://www.securitytransformation.org/gc_publications.php

Shelley, L. I. (2002). The Nexus of Organized International Criminals and Terrorism. International Annals of Criminology, 1-6.

Shelley, L. I. & Picarelli, J. T. (2002). Methods Not Motives: Implications of the Convergence of International Organized Crime and Terrorism. *Police Practice and Research*, *3*(*4*): 305-318.

Shelley, L. I. (2005). The Unholy Trinity: Transnational Crime, Corruption, and Terrorism. *Brown Journal of World Affairs*, *11(2)*: 101-111.

Sparrow, M.K. (1991). The application of network analysis to criminal intelligence: An assessment of the prospects. Soc. Netw. 13, 251–274.

Trifunovic, D. (2007). Special Report: Terrorism and Organized Crime in South-Eastern Europe: The Case of Bosnia-Herzegovina, Sanzak, and Kosovo. *Defence & Foreign Affairs/International Strategic Studies Association*.

Van Duyne, P. C. (2003). Organizing cigarette smuggling and policy making, ending up in smoke, Crime, Law and Social Change, 39, 285-317.

Vasic, M. (2009). No one can fight alone against organized crime: Interview with the Minister of Interior of the Republic of Serbia. *Vreme Magazine*

Von Lampe, K. (2006). The cigarette black market in Germany and in the United Kingdom. *Journal of Financial Crime*, 13, 235-54.

Warburto, J. (2007). Driving to anti-corruption utopia without a roadmap. Assessing the theoretical basis for Corruption Prevention Strategies of Global Institutions. Centre for Research and Studies in Sociology, CIES e-working paper N° 31.

Williams, P (2001) 'Transnational Organised Crime' in Arquila, J and Ronfeldt, D (eds.) Networks and Netwars: The future of terror, crime and militancy.

Including female nonmainstream representations in NGO mainstream discourses to enhance peace building

Voula Kyprianou¹, Paul Knepper¹, and Filippos Proedrou²,

¹ School of Law, University of Sheffield, ² Department of Politics and International Relations, City College (Thessaloniki) pkyprianoul@sheffield.ac.uk

Abstract Despite the vast research on wartime violence against women and their experiences and roles during and after the conflict, little is known about women with a wide range of wartime and postwar experiences. Literature tends to present women as peaceful and passive victims of militarism and nationalism. This research aims to explore how international and local institutions illustrate - and to a great extend contribute to - the construction of women's identities and roles during the post-conflict process to enhance peace. It attempts to do that through the study of NGO reports on Bosnian crisis, although this can apply to other postwar cases as well. An analysis of these documents intends to challenge the assumption that NGO efforts to empower postwar women's position and support peace process should exclusively focus on female civilian victims. Then it attempts to explore views of NGO members who work in postconflict zones on whether stories and perceptions of women with more diverse and complex roles should be included into their NGO accounts and how this can be achieved. It also intends to explore views from women with a wide range of activities on being included in NGO peace-building accounts and their potential contribution to institutional efforts to enhance peace.

Keywords: *Wartime violence, peace building, NGO policies, female fighters.*

1 Introduction

Wartime sexual violence against women is often the result of the process of militarization and nationalism. As this process reaffirms patriarchal assumptions of traditional gender roles with men and women being illustrated as protectors and bearers of the nation respectively, hidden under wartime chaos sexual violence against women increases. Sexual violence in conflict has been expressed in different and various forms, such as blackmailing, rape threats, sexual harassment and rape.

Literature tends to present women as passive wartime victims. Women are usually portrayed as victims of sexual assault, refugees or displaced persons during and after the conflict. Their stories focus mainly on missing persons and they are based on their experiences as mothers and/or wives. Female victimization is often intertwined with women's experiences with nurturing and caregiving, and explained in the discursive framework of difference feminism; men are viewed as relatively violent and war is connected to male sexuality and the aggressive nature of gender relationships¹. Women are seen to be more peaceful and capable of conflict resolution due to motherhood.

Nonetheless, women are often illustrated to be preservers of culture and supporters of violence against 'enemy'. Such support is perceived as an attempt to protect their families and nation, indicating their loyalties to their communities². This can be seen, for example, in nationalistic images of mothers sacrificing their sons for the common good. But although literature includes women in political and militaristic discourse, this activism is usually linked to their peaceful nature through their participation in peace movements.

Studies on female ex-combatants in post-conflict Bosnian society and their roles in peace building seem to be displayed within the wider scholarship on ex-fighters' global experiences of demobilization and reintegration in postwar communities. Emphasis is given on local and international efforts to provide ex-fighters with the necessary training and education in order to enter the job market and gain financial stability. Literature often stresses the limitations of DDR programs, especially for female ex-combatants. Women who participated in wartime violence are more likely to be stigmatized because they had gone beyond their traditional societal roles. Many of them do not wish to undertake their traditional roles after war and face difficulties in rebuilding their relationships with their families.

With regards to NGOs, literature is merely constrained on the contribution of the organizations to DDR programs by offering women psychological support and health programs mainly on reproductive issues. It highlights NGO financial and technical support in encouraging ex-combatants to engage in development projects. Literature on Bosnia shows that demobilization and reintegration process has encouraged NGOs to undertake truth-telling initiatives; some NGOs have given the opportunity to the public to ask questions to demobilized soldiers and hear the 'other side'.

While literature on DDR programs offers insight on ex-combatants' contribution to reconstructing post-conflict societies, little has been said about female ex-combatants and their potential contribution to peace building. Although reconstruction of post-conflict societies is an inextricable part of peace building, this however should not necessarily be seen as willingness on the behalf of ex-combatants to participate in multi-ethnic activities and cross community cooperation. Projects to help ex-fighters enter the job market are essential in social rehabilitation but cannot guarantee cooperation among all ethnic groups to eliminate mutual ethnic suspicion and prejudice against ex-combatants and support peace building.

¹ Goldstein, J. S. (2003), *War and Gender: How Gender Shapes the War System and Vice Versa*, Cambridge University Press, New York, pp. 42-47

² Ruddick, S. (1991), From Maternal Thinking To Peace Politics, in Browning C. E. and Coultrap - McQuin, S. (eds.), *Explorations in Feminist Ethics: Theory and Practice*, Indiana University Press, Indiana, p. 147

Literature also inclines to examine how DDR program could improve their services; should institutions initiate these programs or ex-combatants themselves? In addition, literature tends to analyze the financial, technical and psychological contribution of NGOs in encouraging ex-combatants to actively participate in post-war initiatives to reconstruct their societies, marginalizing other potential NGO contributions to support rebuilding of postwar societies and peace building. My research stresses the need to move beyond DDR discourse when studying about potential roles of female ex-combatants in peace-building process and to explore NGO potential contribution to multiethnic and cross communities activities of female civilians and ex-fighters through sharing their wartime and postwar experiences and perceptions.

This research aims to analyze how organizations present and construct women's post-war identities and roles in an attempt to strengthen the peace process and deal with gender inequalities. It also explores whether and how women with main or secondary combat (and noncombat) roles could or should be included in NGO peace building accounts and efforts from peace-builders', NATO and NGO members' point of view as well as from women with multiple and complex wartime and postwar roles. This thesis attempts to (a) offer a general background of wartime violence theories as explored in the existing literature, (b) localize these theories and display them into the context of the Bosnian conflict - although they can apply in other cases of constructing post-war communities, (c) examine the role of feminine traits and female victimization in forging institutional policies and the effects (and limitations) of those policies in constructing female post-war identities and sustaining peace. (d) collect and analyze views of NGO members who work in post-war societies on including wartime and post-war non mainstream representations into their accounts (e) explore views of women with a wide range of wartime and postwar experiences on including their wartime and postwar testimonies in NGO peace-building accounts and their potential role in peace process outside the DDR framework and (f) make some suggestions about including women who do not fit the traditional norms into NGO policies for gender equality and peace building.

2 Literature review

The above overview draws on a wide range of literature on female wartime representations and sexual violence. Conflict has proved to render women more dependent on others; humiliation of the 'enemy' appears to be a major motive for sexual abuse against women. This does not however underline personal relations and motives as well as feelings of misogyny when examining female wartime representations. For instance, increasing violence in post-communist conflicted Yugoslavia was partly due to suppressed misogyny feelings during communism (Nikolić-Ristanović, 1996; Nikolić-Ristanović, 2000). Misogyny as a factor of violence against women has also been explained on the grounds of gendered victimization, with women being the main victims of aggression. Such violence is often attributed to male fears of social connection and dependence (Goldstein, 2001), efforts to bolster masculinity (Turpin & Lorentzeu, 1996), or they are placed within the context of unequal gender relations (Enloe, 1990).

Although these studies seem to offer different interpretations on wartime sexual abuse - usually against women - they often present similar argumentation and enhance each other. They are placed within the context of patriarchy, masculine ideals and militaristic ideologies. For example, Nikolić-Ristanović's emphasis on the role of patriarchal attitudes about identity in committing sexual abuses could be accompanied with Littlewood's assumption that rape is the result of increasing anxiety and feelings of autonomy – as can be concluded in his article *Military Rape*. They are both based on a patriarchal system, where masculine ideals prevail and justify sexual violence. Littlewood goes further claiming that protection is often exchanged with sexual exploitation (Littlewood, 1997). Indeed, analytical studies on militaristic male representations as protectors (Enloe, 2004) and on nationalistic female representations as reproducers (McInnes, 1998) tend to link sexual violence to male action and protection.

Most literature focuses on the collective motives of sex violence against women. Gilboa for instance, refers to directed sexual abuses by army generals to increase soldiers' morale (Gilboa, 2001). This argument has often placed wartime sexuality into the context of governmental efforts to distract the public from corruption (Goldstein, 2001). Fewer studies however take into consideration personal motives; they often incline to outline personal pleasure as a prime factor in actions of wartime sexual abuse (Hagen, 2010). Personal motives of (sexual) abuse during war are often analyzed by those scholars who argue that (sexual) violence is rather a side effect of war due to disruption than planned and organized (Kalyvas, 2002).

Feminist and wartime sexual violence theories shed a light on female civilian victims' wartime experiences and postwar roles in reconstructing their societies and enhancing peace. But these theories say little about female wartime and post-conflict nonmainstream representations. Literature on female ex-combatants and their roles in rebuilding postwar societies focuses on DDR programs, their aims, contribution and limitations in the process of reconstruction. Scholarship on Bosnian female exfighters reflects the wider literature on female ex-combatants in any post-conflict society. Emphasis is paid on DDR contribution to developing sustainable societies through personal responsibility on the behalf of ex-fighters (Colletta, Kostner & Wiederhofer, 1996). These studies often entail a number of DDR cases in various post-conflict societies, describing the different processes, particularities and limitations of DDR programs in a wide range of post-conflict regions.

Some fewer, exceptional studies outline ex-combatants' potentials in DDR programs and their capacities in rebuilding their communities (Rolston, 2007) or highlight ex-combatants' peace-building skills, which have been gained during participation in war (Coulter, Persson & Utas, 2008). They do not however deal with NGO strategies; they limit their analysis on DDR discourse or they fail to explore how these existing skills could be transformed in an attempt to fit the purposes of peace building. Some others seem to be more gendered specific, either examining programs where male ex-combatants participated in an attempt to change masculinity (McNairn, 2005) or by outlining the financial implications of excluding female exfighters from DDR (Zuckerman & Greenberg). But such implications are more directed to community development rather than to peace building per se. Gender and peace building is often analyzed within the framework of gender mainstreaming impact on society's expectations regarding both genders' potential in peace building (International Alert, 2000).

Reports on Bosnia indicate that most Bosnian ex-combatants had not participated in DDR programs (Moratti & Sabic-El-Rayess, 2009). Literature analyzes NGOs contribution to encourage ex-combatants to actively participate in postwar initiatives to reconstruct their societies (but not necessarily peace building), without challenging existing female representations. My research stresses the need to move beyond female representations as portrayed in DDR (and victimization) discourse and to explore NGO potentials in reconstructing more complex and diverse female representations in and for peace building.

Literature therefore provides two main female representations; female civilian victims and female ex-combatants. Studies on female ex-combatants seem to differ from mainstream discourses on female representations of civilian victims, but similarly to mainstream stories, such studies seem to ignore the complexity and diversity of their identities and agencies in and for peace building. My study suggests that institutional efforts to enhance peace building should escape monolithic and stable female representations either as victims or as ex-combatants and explore those identities and agencies from a different perspective. However, I will not attempt to explore exclusively this particular group, but female ex-combatants will be seen as part of a wider group of nonmainstream representations, that is women with multiple and more diverse roles and experiences, who are not included into institutional documents and their potentials in peace building.

Literature provides a wide range of studies on wartime sexual abuse on women. But the proposed research attempts to address an unexplored area, which is how NGOs use this discourse to - intentionally or unintentionally - (re) construct female post-war identities and agencies and to support the peace building. Could or should they include women who diverge from mainstream wartime and feminist discourses? The research therefore intents to address the complexity of including nonmainstream representations in NGO peace building, gendered accounts.

3 Data and methodology

My literature will derive from feminist discourse and wartime rape theories. This theoretical background will be enhanced by a systematic analysis of NGO documents on conflict and gender issues and a study of a wide spectrum of NGO members' views on including nonmainstream female representations into their peace building efforts.

Desk research

I will be exploring NGO documents on their post-war policies, which allow us to study how wartime and post-war female representations are illustrated in NGO peace building accounts to support the peace process. The selection, study, interpretation and comparison of NGO documents seek to explore how feminist accounts and war theories fit into institutional policies and their implications in resolving gender issues and promoting peace.

Sources – (a) NGO documents: reports from Amnesty International (AI), International Museum of Women, Cooperative for Alliance and Relief Everywhere (CARE), Women's Initiative Foundation, The United Methodist Committee on Relief (UMCOR), Center of Nonviolent Action (CNA), Institute for War and Peace Reporting, Human Rights Watch, Association for Rehabilitation of Torture Victims – Centre fro Torture Victims (CTV), and Women For Women.Org. I have obtained institutional reports via online archives of the various organizations, accessed through their websites.

(b) Media releases: press articles from New York Times, which have been published within the period 1996 until 2012 and allow us to examine the image of women at wartime and after war as illustrated in the press.

I have chosen NGO documents and press articles on Bosnian crisis because of the enormous number of available NGO reports on Bosnian crisis. This will enable me to study one case in detail. The case is particular in terms of the intensity of sexual wartime violence abuses reporting, but it can provide insight into issues that apply in every post-war society. Although Bosnian female representations in NGO reports on peace building will be studied in detail, the main focus is on how NGOs that deal with these issues in any post-war society can include non-traditional female representations into their accounts and policies. I therefore use an analytical model that can be generalized and exist in any post-war case.

Semi-structured interviews

I will be attempting to find out about institutional members' perceptions on including women with multiple and diverse wartime and postwar experiences in peace process. Such members include NATO staff, NGO members and peace builders. Their views should not necessarily reflect the strategies of the organization they work for, although I recognize the difficulties in following a different approach from that of their organizations'. I may need to ask permission from the president of the institution in order to talk with members about the policies of the institution. Or I may be constrained to talk only to a certain group of members, leaving out certain sections of people who work for that organization. In addition, some members might not speak their own voice, or they might not be objective because they feel obliged to talk in a particular way.

This however does not mean that institutional members are not an essential source of information; it still helps me gather some information about organizations' views on the matter and/or their own perceptions based on their work experience. In any case, this source allows me to examine the benefits and limitations of including women with a wide range of wartime and postwar experiences in NGO policies to support peace from the organizations' point of view. My interviews will be semi-structured so that important issues are discussed and covered, but interviewees will have the opportunity to talk about other related areas.

Sample - The sample of participants should not be solely constrained to members of institutions in Bosnia. This is because the questions, which are addressing for institutional members aim to bring new insights to institutional policies in post-war conflicts, regardless of the region. They should all however deal with gender inequalities, wartime victims and peace building. Flexibility in who I interview is therefore essential as long as the participants can give information to interview's questions and does not contain systematic biases due to a lack of institutions' variety.

Semi-structured interviews – I will be attempting to conduct interviews to women with a wide range of wartime and postwar experiences and listen to their voices with reference to their understandings about their post-war roles within a society that is constructing peace, and the insights they can bring to NGOs based on their roles and activities during and after conflict in order to enhance the peace process. Such representations include veterans, women involved in incidents related to property, witnesses of incidents, women who provided moral and health support for victims/perpetrators. I recognize potential difficulties in interviewing nonmainstream women, since they may not wish to discuss anything that is traumatic for them or may upset them. For this reason, I have chosen specific questions that take into consideration the sensitivity of the issue. The interviewes could expand their answers only if they wish. The questions are simple and related to their potential contribution in a peace-building society rather than to their past actions.

Trust, confidentiality, anonymity and consent to participate

I will be introducing the topic of my research before the interview. I will ensure the participant about their rights to anonymity and confidentiality, and withdrawal from the interview process anytime. I will inform them about the tape-recording before interviewing and make clear that participation is voluntary and does not impose any risks to participants.

Objective assessment

My research recognizes the need for further study in order to include accounts from women with a wide range of activities. It attempts to answer whether this is feasible or not, its limitations, challenges and advantages when reconstructing post-war societies and enhancing peace, through partialness, recognition of the existence of multiple realities and views and efforts to elevate this diversity in perceptions on the matter. A combination of methods – NGO documents analysis and interviews – as well as sources – NGO reports, press articles, a wide range of members from different institutions – and awareness of their limitations also support my efforts to conduct my research in accordance to my academic principles to achieve objectivity.

4 Empirical analysis

My hypothesis will suggest that sustainable NGO efforts to empower post-war women's position and support peace process should result from the inclusion of a wide and diverse range of female wartime and post-war representations, their stories and perceptions into NGO accounts. Although institutional policies include female fighters in DDR programs with emphasis on skill training and employment, NGOs tend to ignore nonmainstream or more complex female representations as essential actors in enhancing the peace process.

Indeed, female nonmainstream and more complex representations often seem to lack these 'feminine traits' inherited in female civilian victims, such as motherhood, which are considered to be the primary functions of social life which reconstruction depends on. This proposed hypothesis aims to obtain views and perceptions of institutional members and women with various wartime experiences on potential inclusion of more diverse female representations into institutional accounts and policies to promote peace.

Within this overarching aim there is a range of objectives, such as:

- (a) The critical study on how NGO mainstream discourse affects (re) construction of (gendered) identity and agency in peace building through analyzing NGO documents.
- (b) The analysis of how NGOs deal with women who participated in a wide range of activities. How do traditional gender norms affect NGO policies when dealing with such cases? How do they interpret such actions based on the gendered discourse used in their accounts?
- (c) The examination of how NGOs could change their policies including a diverse range of (more complex) female representations without risking the unity of the organization, contradicting their policies, and hindering the peace process. This will be explored based on a wide range of perceptions from people who work/lived in post-war societies.
- (d) The exploration of how if possible NGOs could use female mainstream and nonmainstream representations to challenge existing institutional discourses. For instance, how could existing NGO discourses on victimization differentiate with the inclusion of women with multiple activities/experiences into institutional accounts? Why is this important? How would such inclusion affect concepts of identity in peace building and peace building per se?
- (e) The identification of potential limitations and challenges with the inclusion of diverse, more complex female representations in peace building as explored by institutional members and possible ways of overcoming them.

(f) The analysis of female wartime and postwar nonmainstream representations' potentials in peace building outside the DDR discursive context.

My research findings could provide some essential insights to other similar experiences in other regions. It has implications for all NGOs, which deal with gender issues and conflict.

5 Conclusions

My research's conclusions pertain to the question of including women with multiple roles and activities into NGO peace-building accounts. Such inclusions may result into a more balanced approach on behalf of NGOs when dealing with post-war (gendered) issues and peace building. But what do members of various institutions think about the advantages, limitations and challenges of such changes? What do women with a wide range of wartime (or postwar) activities themselves believe about their personal capabilities and potentials in peace building with NGOs' support? The research on changes in institutional mainstream discourse aims at drawing conclusions on NGOs role and potentials in representation and on improving NGO gendered, postwar policies to strengthen the peace process.

References

Gilboa, D. (2001), Mass Rape: War on Women, 1-7

Jones, A. (2001), Genocide and humanitarian intervention: incorporating the gender variable

Moratti, M. and Sabic-El-Rayess, A. (2009), *Transitional justice and DDR: the case of Bosnia and Herzegovina*, International Centre for transitional justice, 1-39

Connell, R. W. (2005), Masculinities, Polity Press, Cambridge

Colletta, N. J., Kostner, M. and Wiederhofer, I. (1996), *Case studies in war-to-peace transition: the demobilization and reintegration of ex-combatants in Ethiopia, Namibia, and Uganda*, The World Bank, Washington

Coulter, C., Persson, M. and Utas, M. (2008), *Young female fighters in African wars: conflict and its consequences*, Nordiska Afrikainstitutet, Uppsale

Cudd, A. E. (2006) Analyzing oppression, Oxford University Press, New York

Eagleton, M. (2003) A Concise Companion to Feminist Theory, Blackwell, Oxford

Enloe, C. (1990), *Bananas, beaches and bases: making feminist sense of international politics*, University of California Press, Berkeley

Enloe, C. (2004), *The curious feminist: searching for women in a new age of empire*, University of California Press, Berkeley

Galligan, Y., Clavero, S. and Calloni, M. (2007) *Gender politics and democracy in post-socialist Europe*, Barbara Budrich Publishers, Leverkusen

Goldstein, J. S. (2003), *War and gender: how gender shapes the war system and vice versa*, Cambridge University Press, New York

Grant, J. (1993) Fundamental feminism: contesting the core concepts of feminist theory, Routledge, London

Kalyvas, S. (2000), *The logic of violence in civil war*, Cambridge University Press, New York

Keen, D. (2008), Complex emergencies, Polity Press, Cambridge

Lloyd, M. (2005) *Beyond identity politics; feminism, power and politics*, SAGE Publications Ltd., London

McInnes, J. (1998), The end of masculinity, Open University Press, Buckingham

McLaughlin, J. (2003) Feminist social and political theory, Palgrave McMillan, New York

Meyers, D. T. (2002) *Gender in the mirror: cultural imagery and women's agency*, Oxford University Press, New York

Nikolic-Ristanovic, V. (2000), *Women, violence and war: wartime victimization of refugees in the Balkans*, Central European University, Budapest

Tickner, A. J. (1992), Gender in international relations: feminist perspectives on achieving global security, Columbia University Press, New York

Armstrong, M. (2001) Fear of a real planet: sublunary fantasies of gender, sex, and nation, in Currie, G. and Rothenberg, C. (eds), *The (re) visions of the subject; land-scape, ethnoscapes and theoryscapes*, Lexinghton, Oxford, 217-237

Bunch, C. (1991), A global perspective on feminist ethics and diversity, in Cole, E. B. and Coultrap – McQuin, S. (eds), *Explorations in feminist ethics: theory and practice*, Indiana University Press, Indiana, 176-186

Burns, J. (1996), The media as impartial observers or protagonists: conflict reporting or conflict encouragement in Former Yugoslavia, in Paterson, J. R. and Preston, A. (eds), *Bosnia by television*, British Film Institute, London, 92-100

Drakulić, S. (1993) Women and the new democracy in the Former Yugoslavia, in Funk, N. and Mueller, M. (eds), *Gender politics and post-communism; reflections from Eastern Europe and the Former Soviet Union*, Routledge, New York, 123-130

Eisenstein Z. (1993) Eastern European democracies: a problem of unequal equality, in Funk, N. and Mueller, M. (eds), *Gender politics and post-communism; reflections from Eastern Europe and the Former Soviet Union*, Routledge, New York, 303-317

Frazer, N. (2002), After the family wage: a postindustrial thought experiment, in Mui, C. L. and Murphy, J. S. (eds), *Gender struggles: practical approaches to contemporary feminism*, Rowman and Littlefield Publishers INC, Oxford, 69-94

Goldman L. B. (1996), To act without 'isms': women in East Central Europe and Russia, in Turpin J. and Lorentzen L. A. (eds), *The gendered new world order: milita-rism, development, and the environment,* Routledge, New York, 35-50

Hagen, T. K. (2010), The nature of psychosocial consequences of war rape for individuals and communities. *International Journal of Psychological Studies*. 2, 2, 14-25

Hekman, S. (1999) Identity crises: identity, identity politics and beyond, in Hekman, S. (ed), *Feminism, identity and difference*, Frank Cass Publishers, Essex, 3-26

Hirschmann, N. J. (2002), Toward a feminist theory of freedom, in Mui, C. L. and Murphy, J. S. (eds), *Gender struggles: practical approaches to contemporary feminism*, Rowman and Littlefield Publishers INC, Oxford, 47-68

Hofman, A. (2009) 'When we were walking down the road and singing': rural women's memories of socialism in Serbia, in Penn, S. and Massino, J. (eds), *Gender politics and everyday life in state socialist Eastern and Central Europe*, Palgrave Macmillan, New York, 185-197

Kalyvas, S. and Sambanis, N. (2005), Bosnia's civil war: origins and violence dynamics, in Collier, P. and Sambanis, N. (eds), *Understanding civil war: Europe, Central Asia and other regions*, World Bank Publications, Washington, 191-229

Langton, R. (2000) Feminism in epistemology; exclusion and objectification, in Fricker, M. and Hornsby, J. (eds), *The Cambridge companion to feminism in philoso-phy*, Cambridge University Press, Cambridge, 127-145

Littlewood, R. (1997), Military rape. Anthropology Today. 13, 2, 7-16

McNairn, R. (2005), Building capacity to resolve conflict in communities: Oxfam experience in Rwanda, in Sweetman, C. (ed.), *Gender, peace building, and reconstruction*, Oxfam, Oxford, 83-93

Marchand, M. H. and Parpart, J. L. (2003) Exploding the canon: an introduction/conclusion, in Marchand, M. H. and Parpart, J. L. (eds), *Feminism, postmodernism, development*, Routledge, New York, 1-22

Meznarić, S. (1985) Theory and reality: the status of employed women in Yugoslavia, in Wolchik, S. L. and Meyer, A. G. (eds), *Women, state, and party in Eastern Europe*, Duke University Press, Durham, 214-220

Milić, A. (1993) Women and nationalism in Former Yugoslavia in Funk, N. and Mueller, M. (eds), *Gender politics and post-communism; reflections from Eastern Europe and the Former Soviet Union*, Routledge, New York, 109-122

Mui, C. L. and Murphy, J. S. (2002), Introduction, in Mui, C. L. and Murphy, J. S. (eds), *Gender struggles: practical approaches to contemporary feminism*, Rowman and Littlefield Publishers INC, Oxford, 1-46

Nikolic-Ristanovic, V. (1996), War and violence against women, in Turpin, J. and Lorentzeu, L. A. (eds), *The gendered new order: militarism, development and the environment*, Routledge, New York, 195-210

Rolston, B. (2007), Demobilization and reintegration of ex-combatants: the Irish case in international perspective. *Social Legal Studies*, 16, 2, 259-280

Ruddick, S. (1991), From maternal thinking to peace politics, in Browning C. E. and Coultrap - McQuin, S. (eds.), *Explorations in feminist ethics: theory and practice*, Indiana University Press, Indiana, 141-155

Turpin, J. and Lorentzeu, L. A. (1996), Introduction, in Turpin, J. and Lorentzeu, L. A. (eds), *The gendered new order: militarism, development and the environment*, Routledge, New York, 1-12

Wolchik, S. (1989), Women and the state in Eastern Europe and Soviet Union, in Charlton, S. E. M., Everett, J. M. and Standt, K. A. (eds), *Women, the state, and development*, State University of New York Press, Albany, 45-65

Zuckerman, E. and Greenberg, M. (2005), The gender dimensions of post-conflict reconstruction: an analytical framework for policymakers, in Sweetman, C. (ed.), *Gender, peace building, and reconstruction*, Oxfam, Oxford, 70-82

Greek Memorandums of Understanding & Democracy

Alexandros Kyriakidis1

¹PhD Candidate – Politics Deprtament, University of Sheffield,

a.kyriakidis@sheffield.ac.uk

Abstract. The financial crisis of the late-2000s has impacted states around the globe greatly, and perhaps has hit more severely many of the & zone Member States, initiating deep and controversial policy changes. In this research the case study of Greece will be investigated in terms of the political management of the crisis, i.e. mainly its economic adjustment programs (known as Memorandums of Understanding), drafted by the European Central Bank, the European Commission and the International Monetary Fund, and the effect these programs had on the accountability and legitimacy of the Greek state. The research question of the investigation is: What are the effects, if any, of the MoUs implemented in Greece, on the democratic value of the Greek state? A detailed content analysis of core documents is employed in a comparative framework, outlining the contradictions, if any, between Memorandums of Understanding, which aim to accompany funds provided to the Greek state to counter the effects of the crisis, and the Greek constitution, in which most of the democratic principles of legitimacy and accountability are established (in a state of the civil law tradition). It is concluded that, perhaps contrary to judicial precedent, the MoUs in Greece stand at complete odds with Greece's democratic principles of accountability and legitimacy, and constitute a clear violation of the self-determination of the Greek state. The results of the investigation are introduced into the wider framework of the relationship between democracy and structural adjustment programs, combined with the use of said programs during times of crises as vehicles to significantly alter the status quo with minimum resistance.

Keywords: Political Economy, zone Crisis, Crisis Management, Democracy, Greece

Abbreviations

Abbreviation	Full Term
€MS	Eurozone Member States
ABS ¹	Asset-Backed Securities
Alt-A ²	Alternative A-paper
CDOs ³	Collateral Debt Obligations
CG	Constitution of Greece
CoS	Council of the State
EC	European Commission
ECB	European Central Bank
EU	European Union
GDP	Gross Domestic Product
IMF	International Monetary Fund
MBS ⁴	Mortgage-Backed Securities
MoUs	Memorandums of Understanding

1 Introduction

1.1 Framework of the Research

The general framework of this investigation is structural adjustment programs, implemented by International Organizations (IOs) in financially, and oftentimes politically, troubled countries in return for low-interest loans (usually aimed towards development), and the impact of said on the democratic value of these countries. While said programs have been drafted mainly between the International Monetary Fund (IMF) in developing countries, the financial crisis of the late 2000s was the cause for an abrupt change of that, marking the first time that the IMF implemented the aforementioned type of programs in developed Western democratic states, and in doing so de facto directly interfering in the affairs of the & zone / European Union (EU).

Not only that, but the crisis also marked the beginning point of EU institutions initiating similar programs with the IMF, adopted by & zone Member States (&MS) that suffer from a severe credit squeeze in turn for financial assistance. The European

¹ Bonds that are backed by financial assets consisting "of receivables loans, such as credit card receivables, auto loans, manufactured-housing contracts and home-equity loans." (Investinginbonds.com 2010).

² Mortgages that are neither prime (A-paper) nor sub-prime (maximum risk), but lie in the middle (Investopedia 2013).

³ "An investment-grade security which is backed by a pool of bonds, loans and other assets." (Investopedia , 2012).

⁴ "Mortgage-backed securities (MBS) are debt obligations that represent claims to the cash flows from pools of mortgage loans, most commonly on residential property." (US Securities and Exchange Commission, 2012).

Commission (EC), the European Central Bank (ECB) and the IMF, what is commonly referred to as the Troika [European Union (c) 2013], worked closely in cooperation (although drafting separate program requirements and providing individually different forms and amounts of financial assistance) to draft plans of structural adjustment programs, known as the Memorandums of Understanding (MoUs)⁵, that were to be applied by troubled \in MS in turn for financial assistance.

The research question of this investigation is set as follows: What are the effects, if any, of the Greek Memorandums of Understanding, on the democratic value of the Greek state? The hypothesis of the research is that: There have been negative effects of the Greek Memorandums of Understanding, on the democratic value of the Greek state. The investigation will conduct an in depth, comparative analysis between the MoUs, and the Constitution of Greece (CG). The constitution, at least for a state that belongs to a civil law tradition such as Greece, is considered the foremost important document that contains all the democratic principles and processes of a state, most fundamentally safeguarding in a number of different ways the principles of democratic legitimacy and accountability. Any contradictions found will yield a positive answer to the hypothesis, suggesting that the MoUs have rendered Greece less democratic.

Greece as a case study presents an intriguing case. As evident from Appendix A of this paper, Greece has received the most amount of financial assistance (\in 140.8 bln - almost double from Ireland that comes in second place) and the summation of all pages of IMF/EU drafted MoUs amounts to the amazing 1021 pages, more than 300 above Portugal, which comes in second place (Appendix A). Notwithstanding the tremendous size of the MoUs, and as will be argued in this paper, the intricacy, level of detail and extent to which the MoUs affect policy in almost all areas of policy that a government of a state is responsible, are incredibly deep.

It is a fact that some (albeit not all) of the MoUs have been ratified by the Greek Parliament and have thus been transformed into law, usually under a somewhat weak majority, with many additional others that deal with the specific aspects of the MoUs, such as for example Law No 3986/11 which, among others, establishes the Hellenic Republic Asset Development Fund (HRDAF). Moreover, it is also the case that there has also been judicial precedence over the general constitutionality of the MoUs. The Council of the State (CoS), the supreme constitutional court of Greece [Council of the State (a) 2012] heard and deliberated, under plenary 668/12, on whether four Ministerial Decisions (MDs) / circulars, which provided frameworks and specific guidelines on applying the Laws that enforce the MoUs, were unconstitutional. The CoS ruled in favor the MDs / circulars, based on a number of reasons, evaluating that the above did not violate the CG [Council of the State (b) 2012].

Hence, it could be argued that since the MoUs (at least the ones that were voted on) underwent voting in the Parliament, they are democratic. After all, the judiciary as

⁵ The following fall under the term MoU: Memorandum of Economic & Financial Policies, Memorandums of Understanding on Specific Economic Policy Conditionality. The Technical Memorandums of Understanding, although included with the above, are not examined in this essay, as they contain mostly technical and economic information as well as definitions of terms such as deficit, GDP, State enterprises, etc.

well as the executive and the legislative, all seemed to concur that these agreements were not unconstitutional. However, violation of a constitution, or violation of democratic value, is not precluded by a reached agreement of all three separate powers. It is not uncommon that democracies fail to operate correctly. Therefore, despite judicial precedence and agreement of all three powers, a more in-depth political analysis is needed, with a detailed comparative structure, to evaluate whether the MoUs do or do not violate the democracie value of the Greek state as provided by the CG.

Aside from the above, there are a number of MoUs that were signed and did not transform into law. The aforementioned international agreements that have not been ratified by the Greek Parliament, constitute the majority of the MoUs and have been signed only by the executive branch of the Greek government. Albeit having no legal ground within the Greek state, these international agreements, which go at great lengths to define every single aspect of almost all policy areas within Greece, take shape and form in the Greek state within isolated blocks of legislation addressing different aspects of them. However, prior to conducting the investigation, it would be advisable to provide a brief overview of the timeline of the events of the crisis, with a particular focus on Greece, to sufficiently establish the general conditions during the time of the research.

1.2 A timeline of the crisis

Financial crises have differed greatly across time, starting as early as the Tulip-mania bubble of 1637 (Sinclair 2010, 93-4). This investigation relates to the current, ongoing crisis, the origins of which can be traced as far back as the credit boom of stocks relating to dot.com bubble and the ensuing stock market frenzy during the 2000s, which was arguably preceding boom, after which the crisis followed (Sinclair 2010, 93 & 100). Although many starting points have been argued in relation to the crisis, it is now almost certain that the core problem that led into the eventual credit crunch⁶ was subprime mortgages and the decision to withdraw investment from them (stop purchasing related securities), starting as early as the winter of 2007.

During the 2000s boom of the US housing market, the practices of predatory lending⁷ were used in order to produce higher profits. Predatory lending led conditionality for mortgages to be severely relaxed, and to Alt-A and subprime mortgages (high-risk products due to low conditionality) being provided as prime mortgages. In the USA alone, the two aforementioned types of mortgages accounted for close to 50% of mortgages for families (27 million), valued at more than \$4 trillion (Wallison 2010, 8). After the housing bubble that had been growing since the late 2000s bursts in 2006, MBS and CDOs values decline drastically.

⁶ A credit crunch or a credit squeeze is a shortage of credit in whatever form, but usually in the form of loans, in a "situation in which the supply of credit is restricted below the range usually identified with prevailing market interest rates and the profitability of investment projects" (Council of Economic Advisors 1991).

⁷ "Predatory lending typically involves imposing unfair and abusive loan terms on borrower [...] ...(It) can be detrimental to consumers and increases the financial and reputation risk for financial institutions" (Office of the Inspector General 2006).

Subprime mortgage products became toxic and their price skyrocketed along with the repayment costs for the individuals who had the mortgages. This created many defaults from people who could not make their mortgage payments, many of whom were not credible enough to be given a mortgage loan, but were given one anyway. These defaults, while leaving people homeless and driving foreclosures way higher than at the Great Depression, also drove real-estate prices down, as more and more mortgages defaulted. In turn, and with MBS heavily traded by that point, the ABS market collapsed. Due to the above, the uncertainty regarding Collateral Debt Obligations (CDOs) (i.e. which financial products were safe and which weren't) led to a decrease in confidence between banks, which in turn led to the restriction of credit "in order to compensate for the uncertainty surrounding possible default" (Martins 2010, 3). The aforementioned conditions, coupled with the generally accepted practice of MTM accounting, led to failure of large financial institutions, which in turn led to a credit squeeze on inter-bank loans, and, correspondingly, in bank loans to individuals (Craigwell et. al. 2011, 793 & Wallison 2010, 8-9).

By March 2008 and onwards, the financial services sector was terminally overrun, with large investment banks, such as Bear Stearns, along with government sponsored enterprises, such as Freddie Mac, and other large financial services firms, such as Lehman Brothers and Merill Lynch, filling for bankruptcy and/or being bailed-out, sold or placed under the control of the USA government (Wallison 2010, 4-7). The most astonishing case of the crisis in the USA, Lehman Brothers (a financial services and investment banking corporation), files for bankruptcy under Chapter 11 of the USA Bankruptcy Code under a massive debt amounting to \$613 billion (stemming from its overexposure to the sub-primed mosrtgage securities), the largest amount over which a bankruptcy has ever been filed in the USA (Mamudi 2008). The domino effect across the banking sector is inevitable, eventually spreading to the financial systems of entire countries across the Atlantic. The crisis was transferred to the &MS in the form of a severe credit squeeze, with Greece being the first victim. Starting from Greece in 2009 and spreading across the \in MS, the &zone provided financial support for Greece, Ireland, Portugal and Spain (the latter only through the ECB) all through 2010-12 (Mead & Blight 2013).

In Greece, it was not until the election of George Papandreou, from the center-left party PASOK, that the crisis really started, spreading quickly to countries with large public debts within the common currency. Dissatisfied with the incumbent Prime Minister (PM) Kostas Karamanlis, of the center-right party New Democracy, mainly due to the worsening economic condition, the people voted for Mr. Papandreou during the October 2009 general national elections, under the promise of improvement of the country's financial situation. Mr Papandreou himself numerous times repeated the, now anecdotal, phrase "There is money" (Greek: $\Lambda\epsilon\phi\tau\dot{\alpha}\,\Upsilon\pi\dot{\alpha}\rho\chio\nu\nu$) (topetopeopapagalos 2010). Under the aforementioned slogan, PASOK won with an approximately 10 point lead, securing 160/300 seats in Parliament (a well secured majority to enact any legislation) (Greek Ministry of the Interior 2009). However, right after assuming office in late 2009, Mr Papandreou admitted that the previous government hugely underestimated the annual budget deficit of the country, and announced that it will be almost double its former value, reaching more than 12% of the national GDP.

In December 2009, two out of the three major CRAs that are officially authorized to conduct ratings by the US SEC, under the license of a "Nationally Recognized

Statistical Rating Organization - NRSRO" (U.S. SEC 1994, 1 & 8-9), Fitch and Standard & Poor's gave Greece a low rating, resulting in a major credit squeeze and inability of the country to roll-over its ever growing and already sizable debt. Greece's debt rose to 113% of the GDP, while at the same time Ireland, amid efforts to salvage its economy, progressively affected by the continuous liquidity assistance provided to the banking sector, preemptively imposed austerity measures and raised the age for retirement eligibility from 65 to 66 years of age (Encyclopedia Britannica 2013).

The third \in MS to be dragged into the crisis is Spain in early February 2010, with a budget deficit of 11% and a national debt of 75% of GDP at the time. José Luis Rodríguez Zapatero, the Spanish PM was quick to adopt austerity measures to prevent a financial meltdown, providing assurances that Spain is different from Greece [Moya (b) 2010]. At the same time, the Euro took a severe hit as the EU seemed unable to agree on a plan for salvaging the Greek economy [Kollewe (b) 2010]. Assurances by the Greek PM that the country can rebound fall into the void (Wood 2010), and by mid-Arpil 2010, Greece officially requested the assistance of the IMF in cooperation with the &MS to salvage the economic situation that the country was in, through drafting a full and complete package that eventually took the form of Memorandums of Understanding (MoU) [Moya (a) 2010]. Ireland experienced similar symptoms, being downgraded constantly by CRAs [Kollewe (a) 2010], while, by May 2010, the austerity in Greece begun to hit the real economy severely, and desperate civilians amounting to thousands clashed with police in the streets of Athens outside the Greek Parliament (theguardian.co.uk 2010) in protest of the austerity imposed. Throughout 2011, both Ireland and Portugal followed suit, requesting financial assistance from the EU/IMF ad hoc cooperation and adopting, in return, severe austerity measures as by these organizations. During the fall of 2011 Italy was swept into the crisis as well and was downgraded by the CRAs. In early 2012, political turmoil swept across the & zone while Greece adopted additional austerity measures, introducing Private Sector Involvement (PSI) to further avoid bankruptcy and the alleged exit from the & zone, (what became known as Grexit).

During late-2012 and early 2013, Cyprus entered the crisis, largely due to its overexposure to Greek bonds, which they themselves had undergone a major value haircut under the Private Sector Initiative (Xarhs123, 2012 & Ryssdal 2012 & AlJazeera 2012). In June 25th 2012, Cyprus officially requested assistance from the IMF, ECB and EC collaboration (AlJazeera 2012) and the package deal was decided almost a year later in March 2013, involving, for the first time ever, a 'bail-in,' i.e. a haircut of deposits to alleviate the overwhelming debt of the banks the deposits are in, along with other measures such as immediate and abrupt restrictions imposed on the banking sector and a full Memorandum of Understanding [The Tribune (a), (b), (c)]. In this investigation, the Greek MoUs will be evaluated against the GC in an attempt to evaluate the effect of the structural programs on the democratic principles of accountability and legitimacy that lie with the CG.

2 Literature Review

2.1 Democracy

Democracy ($\delta\eta\mu\omega\kappa\rho\alpha\tau(\alpha)$) was created and developed in Ancient Greece, predominantly in the city-state of Athens. "The development of democracy in Athens has been a central source of inspiration for modern political thought. [...] Its political ideals – equality among citizens, liberty, respect for the law and justice – have influence political thinking in the West" (Held 2006, 13). Robinson (2007) provides an excellent and condensed description of this political regime at that time: "As it existed in the fifth and fourth centuries, democracy meant that the demos (the people) were sovereign in the deliberations of state. A popular assembly, to which all citizens were invited, met regularly and provided a forum for debating and voting on the most important matters. Representative councils typically prepare in advance the agenda for the assembly meetings. Popular courts, with ordinary citizens serving as jurors, tried legal cases, and administrative officials (such as generals, treasurers, etc) were either elected or chose by lot for relatively brief terms, usually one year. Officials were held to account after their terms of office as a check on corruption. [...] Underlying the development of these institutions were the ideals of freedom and equality" (3).

Democracy during that time was based on three main pillars. The main two, according to Aristotle's Politics were liberty ($\epsilon\lambda\epsilon\nu\theta\epsilon\rho(\alpha)$) and equality ($\iota\sigma\delta\tau\eta\varsigma$), both of which oftentimes overlap in meaning and essence, "...and since the people are in the majority, and the opinion of the majority is decisive, a government must necessarily be a democracy" (Hansen 2007, 171-2). A third pillar was the severe dedication and devotion to the common political life and to the civic duty, both of which were considered an obligation of the utmost importance for every Athenian citizen (Held 2006, 14).

It is worth noting that democracy, in its direct, pure form (as exercised in Ancient Greece for example), "...was poorly regarded by all the Greek philosophers and historian whose writings have survived, including Plato, Aristotle and Thucydides. They depicted it as government by the ignorant or government by the poor" (Birch 2993, 45). For example "...Plato came ever more to the view that political control must be placed in the hands of a minority" (Held 2006, 24-5), and Aristotle, due to the prerequisite of "strict political equality" for democracy was extremely reserved about it (Held 2006, 16-7).

In the modern, Western world "...democracy denotes both a set of political institutions and a set of political ideals. [...] In liberal democratic thought, democracy, liberty and equality form a triad and are often described as the three points of a triangle" (Hansen 2007, 171). Specifying this triad, Beetham (1999) provides a pyramid with three preconditions for the existence of a democratic society: "Open & accountable government", "free & fair elections", and "civil & political rights" (156). In the pyramid, equality is ensured via adherence to the civil and political liberties, and also via free elections, while liberty is ensured through open and accountable government. In other words, three principles are indicative of the existence of a democratic relationship "between governing and governed" (Weir & Beetham 1999, 9):

• Principle of "authorization" via election of representatives, also termed as legitimacy • Principle of "accountability" directly to the people and indirectly regarding actions of civil

servants, etc.

• Principle of "responsiveness" to the will of the people.

Finally, Broome (2012) raises similar concerns arguing that "In this sense, crisis can lead to structural changes in dominant ideas by introducing new 'knowledge regimes' at both the national and global level" (9). The question however still remains of whether these changes, regardless of cause, are made democratically. Brasset et.al. (2011), within the aforementioned framework, presents six indicators of democratic application of governance during crises: "(1) *Transparency*, i.e. the extent to which all interested observers are in a position to inform themselves fully on the core questions and tradeoffs under consideration; (2) *Openness to direct participation;* (3) *Quality of discourse* or the extent to which the legislative debate is legitimized through evidence of ethical reflection and contested practice as well as administrative norms; (4) *Representation;* (5) *Effectiveness;* and (6) *Fairness*" (7).

From the above, it is evident that democracy is based mainly on three interconnected concepts: liberty or freedom, accountability and legitimacy.

2.1.1 Accountability

In regards to accountability, it has been argued that it is of the same normative gravity as legitimacy, bridging the middle ground between input and output democracy. "It underlines the need for transparent decision-making, legal oversight and good administrative procedures without necessarily imposing participatory requirements that would endanger the efficiency of decision making" (Neyer 2010, 907). Accountability can be generally defined as "being answerable to somebody else, to being obliged to explain and justify (in)action – how mandates and contracts have been dealt with, how authority and resources have been applied, and with what results" (Olsen 2013, 3-4). Given the above democracy can indeed be characterized as "… a system of governance in which rulers are held accountable for their actions in the public realm by citizens, acting indirectly through the competition and cooperation of their elected representatives" (Huller 2012, 252-3). This can be achieved via two kinds of accountability (Huller 2012, 252-3):

- "vertical accountability," i.e. citizens holding rulers accountable (elections) and
- "horizontal accountability," i.e. competition between the different representatives

Democracy's main normative standards are that both the decision-makers and the policy processes should be subject to the demos, that misuse of power by decision-makers is avoided, that decision-makers and decisions can always be altered, and that political life is conducted in the public so as to inform citizens of the quality of the decision-makers, to provide assistance towards an informed overall decision (Huller 2012, 254-5). Taking the above under consideration, accountability implies non- or limited - majoritarian institutions, which must conduct themselves according to specific rules and a certain ethos. To ensure that this is the case, appropriate accountability needs to be exercised by informed citizenry, with publicity and transparency hav-

ing a major role in the formation of decision by the populous (Olsen 2013, 23 and Schmidt 2013, 6-7).

2.1.2 Legitimacy

In regards to legitimacy, Birch (1993) provides an indirect definition as "the quality of ascribed entitlement to exercise that (political) power" (32), but mainly resorts to Weber's distinction between three different types of legitimacy (33-4):

Type of Legitima- cy	Overview					
Traditional	Compliance is ensured via direct loyalty to the leader, such as					
	the tribal chief, the king, etc.					
Charismatic	Compliance is ensured through faith to the charisma of a specif-					
	e leader and their qualities, such as Charles De Gaulle for ex-					
	mple.					
Legal-Rational	Compliance is ensured through "general acceptance of the pro-					
	cedures by which these orders and laws are produced" (Birch					
	1993, 34).					

Table 1: Extended,	Specified Legitimacy Types
(Birch 1993,	33-4)

Under the current regime, and given that Greece is a modern, Western, liberal democracy, the legitimacy of the entire political system belongs to the third type as referenced above, i.e. Legal-Rational legitimacy. According to the CG, and especially Articles 1 and 51, it is evident that the political system of Greece is the "constitutional rule of law … (which is) the distinctive liberal-democratic mode of legality" (Beetham & Lord 1998, 9).

Notwithstanding the above, the CG involves all other criteria of Beetham & Lord (1998) that exist in a liberal-democratic state, i.e political authority is achieved via popular sovereignty, the government has the protection of rights and security of its citizens as the primary goal, and elections provide consent by the ruled to the rulers (9). Given the above, a more in-depth analysis of the Legal-Rational model of Legitimacy should be provided. Follesdal (2006) specifies and sub-divides the Legal-Rational type into the following categories:

Table 2: Sub-Categories of Legal-Rational Legitimacy

(Follesdal 2006, 445-50 and Guastaferro & Moschella 2012, 202 and Beetham & Lord

1998, 3-21 and Menendez 2009, 283-4)

Sub-division	Overview
Adjudicative	The judiciary and the decisions of the courts should aim for
	equality among citizens.
Social	Citizens are predisposed and also do abide by laws and author-

	ity.					
Procedural	All citizens have the right to participate in the decision-making					
	and also deliberative policy making processes.					
Normative	Rules & laws and decisions are justified to those who must					
	abide by them.					
Participatory	Include citizens and experts in the decision making process,					
	hrough deliberative and other governance structures.					
Democratic	In the representative democratic system rulers are held ac-					
	countable for their actions, via selection from the people among					
	competing candidates.					
Actual Consent	Citizens decide who is to rule them.					
Output	Achieve policy results that benefit all.					

Of the above sub-categories, it can be argued that participatory, legal-rational legitimacy, also known as the democratic deliberative process, is probably the most important of all. Menendez (2009) argues that "the core of democratic legitimacy is procedural legitimacy, or what is the same, the right of all citizens to participate in the deliberation and decision-making stages of the law-making process. Democratic politics assumes that decisions must be taken, but also that legitimacy is closely tighten to the deliberation which precedes actual decision making (283).

In addition to the above, legitimacy can also be subcategorized according to the objects in regards to which it is examined, i.e. according to (all below from Follesdal 2006, 450-1 and Wimmel 2009, 188-9):

- political decisions,
- political authorities,
- public institutions,
- regime / political order system (along with regime ideals),
- political community (who should be in and out of the community boundaries).

Lastly, legitimacy can have three different variables, according to the stage of the policy making procedure: **Input or Participation**, i.e. how citizens can, directly and indirectly, influence policy or actors making policy, and **Output or Results**, i.e. the evaluation of the results of policy making in relation to the citizens and their preferences / needs (Wimmel 2009, 190-1).

In terms of **output legitimacy**, the concept refers to "...the degree to which the substance of the decision may be said to promote collective interests in a manner compatible with the democratic goals of equal concern and respect" (Bellamy 2010, 3), and aims at resolving problems created in a society in the "interests of the constituency" (Bellamy 2010, 7 and Schmidt 2013, 4 and Hobolt 2012, 93). Prescriptive supporters of output legitimacy (normatively suggest that a polity should be governed by this legitimacy type) do not necessarily deny the lack of citizen input to the decision-making processes of the EU (Schmidt 2013, 5), but rather argue that, whether by design or effect, said input can and should be sacrificed to achieve the desired output (to be produced by non-majoritarian formations). It is supported that increased input might damage policy outcomes leaving the society in a disadvantage position, and thus produce the opposite result of what it aspires to (Bellamy 2010, 3).

In regards to the institutions that conduct policy and produce the output, there are certain preconditions that, if followed, can improve the output legitimacy, and therefore output democracy. Organizations can have common political aims, can be transparent and simple, can be effective and efficient in producing the desired policy outcomes, including producing effective deliberation procedures, can provide for the appropriate social education of common political traditions, can reinforce future citizen compliance through production of stably and fair outcomes and resolving the free rider problem by reporting or imposing sanctions to individuals who take advantage of the outcomes of societal and political deliberation without bearing the appropriate responsibilities, can ensure an incentive based structure, and can implement clear separation of powers with checks and balances (Follesdal 2006, 457-8 and Olsen 2013, 15).

In terms of **input legitimacy**, the concept refers "...to the democratic character of the decision procedure, and in particular the right of all citizens to participate on an equal basis in political decision-making," (Bellamy 2010, 3 and Cheneval & Schimmelfennig 2013, 339 and Schmidt 2013, 4 and Hobolt 2012, 93). Bellamy (2010) also suggests that input democracy is necessary to compensate for mistakes and disagreements occurring between decision-makers (5). In other words, "what counts is being included in the process, having their views treated on a par with everyone else's, and not feeling permanently excluded from consideration. If a core function of democracy is to allow necessary binding collective decisions to be made despite valid disagreements and uncertainty about their potential effects, these qualities seem vital. In addition, they promote mutual respect and reciprocity among citizens" (Bellamy 2010, 5).

3 Research Methods & Design

The research question of this investigation is set as follows: *What are the effects, if any, of the Greek Memorandums of Understanding, on the democratic value of the Greek state*? The hypothesis of the research is that: *There have been negative effects of the Greek Memorandums of Understanding, on the democratic value of the Greek state.* Greece is chosen as a case study between all other seven EUMS that received assistance by the IMF/EU ad hoc cooperation (within the Euroarea: Cyprus, Ireland, Portugal & Spain, outside the Euroarea: Latvia, Romania and Hungary), based on the following three factors, that are presented in detail in Appendix A:

- Amount of money received for financial assistance: & 140.8 bln, almost double from Ireland, which is in second place, with : & 85 bln
- Total amount of MoU pages: 827 pages in total, almost 200 pages more than Portugal's 607 pages in total, which comes in second place.
- Invasiveness and degree of detail in policies enforced via the MoUs: Greek MoUs cover almost every aspect of policy making (health, social security, public employees, state-owened enterprises, fiscal policy, budgetary policy, banking sector reform, taxes, informational infrastructure, education, legislation, judiciary reforms, and others), as will be demonstrated in the analysis section of this investigation, while most of the other states' MoUs refer to

specific, narrow economic prescriptions and targets, with some minimal fiscal consolidation measures. Cyprus comes in second place, with an invasive, but incomparable to Greece, structural adjustment program

Aside from the above, the fact that Greece has always constantly been in the foreground of the Eurozone crisis, in almost all media, as well as the fact that it has admittedly been subjected to possibly the harshest and most detailed structural adjustment program, cannot be ignored. The aforementioned observations can serve as a generic fourth reason for attempting to determine the effect of MoUs on democracy in the case study of Greece. Following below is the detailed arrow diagram of the investigation:

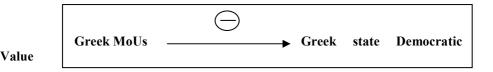


Fig. 1: Arrow Diagram of the Research

The Independent Variable (IV) is the Greek MoUs, operationalized as:

- The seven Memorandums of Economic and Financial Policies, Technical Memorandums of Understanding, and Memorandums of Understanding on Specific Economic Policy Conditionality, that were signed between Greece and the IMF,
- The three Memorandums of Understanding between Greece and the EC.

Date	EU/IMF	Purpose					
03-05-2010	EU	1 st Economic Adjustment Program for Greece -					
		To outline the Greek financial and economic					
		policies "that the Greek Government and the					
		Bank of Grece, respectively, will implement					
		during the remainder of 2010 and in the period					
		2011-2013" [European Commission 2010, 37]					
06-08-2010	IMF	Update to the MoU of 03-05-2010 [IMF (a)					
		2010, 1]					
08-12-2010	IMF	Update to the MoUs of 03-05-2010, 06-08-					
		2010 [IMF (b) 2010, 1]					
28-02-2011	IMF	Update to the MoUs of 03-05-2010, 06-08-					
		010, 08-12-2010 [IMF (c) 2011, 1]					
04-07-2011	IMF	Update to the MoUs of 03-05-2010, 06-08-					
		2010, 08-12-2010, 28-02-2011 [IMF (d) 2011,					
		1]					
30-11-2011	IMF	Update to the MoUs of 03-05-2010, 06-08-					
		2010, 08-12-2010, 28-02-2011, 04-07-2011					
		[IMF (e) 2011, 1]					

Table 3: Greek	k MoUs: Dates,	Organizations	and Purpose
----------------	----------------	---------------	-------------

01-03-2012	EU	PSI rules and agreement for implantation EFSF loan conditions and guidelines Conditionality and prepare the ground for the 2 nd Economic Adjustment Program for Greece				
		[European Commission (b) 2012, 1-7]				
09-03-2012	IMF	To outline the financial and economic policies "that the Greek Government and the Bank of				
		Greece will implement during 2012-2015" [IMF (f) 2012, 1]				
13-12-2012	EU	2 nd Economic Adjustment Program for Greece – To outline the financial and economic poli- cies "that Greece will implement during the remaineder of 2012 and in the period 2013-16" [European Commission (a) 2012, 141]				

There seem to be no Intervening (IV) or Conditional (CV) Variables in this research. The Dependent Variable (DV) is the democratic value, or lack thereof, of the Greek state under and after the MoUs. In regards to operationalizing the dependent variable, a more detailed analysis is in order. It is clear that Greece's democratic value will be investigated through the CG. However, how is democratic value operationalized?

As obvious from the Literature Review section, defining or even operationalizing democracy is a task that researchers have fought with over the centuries, almost for over 2500 years. This research aims at evaluating pragmatically and realistically the effect of the MoUs on the democratic value of the Greek state, without necessarily engaging with the theoretical aspects of democratic theory. Towards this direction, David Beetham first innovatively introduced the democratic audit, meant to assess the democratic value of the United Kingdom.

Beetham (1994 & 1999) manages to side-step the theoretical, definitional problem of democracy, by creating a set of indices for a democratic audit. As Saward (1994) suggests it is possible to "define democracy according to certain basic principles" (7), without having to engage with the 2500 years of political philosophy that the concept carries with it. Hence Beetham (1994), beginning from his basic three democratic principles, as outline in Fig.1, produces thirty indices for the democratic audit of a state following the tradition originating with Robert Dahl (Lord 2008, 1). However, despite the usefulness of his model, Beetham (1994) has predicated it entirely on certain premises that make it inappropriate for this investigation. The model's focus on evaluating inner politics (in contrast to effects from outer-state imposed policy initiatives), its extended number of indices (thirty) that only seem to deal with somewhat specific but multifaceted issues of democracy (such as independence between executive and judiciary, freedom of the press, etc), and its predication upon the fact that "the capacity of political authorities is not separable from their moral standing among those whose cooperation is required for them to achieve their purposes" (Beetham & Lord 1998, 10), all make it relatively useless for evaluating the effect of the MoUs on the value of Greek democracy. It is the aim of this investigation to focus on much more fundamental and basic principles of democracy, given the protest and objections raised against the MoUs, in terms of their democratic value,

Given the above, a summary of the all the concepts outlined in the literature review will be provided, re-categorizing and re-shaping the concepts presented by the different scholars, in order to reach to the creation of a functioning instrument for assessing the MoUs' effect on democracy. Presented below are the different reasons for summarizing and/or putting aside the different concepts offered by scholars in relation to the fundamentals of democracy, i.e. accountability and legitimacy.

 Table 4: Re-organization / abolition of fundamental concepts of liberal democracy

Concept	Research relevant categories				
Positive Liberty	No need to be addressed in the current research, since the				
Negative Liberty	MoUs do not, at least directly, deprive citizens mainly of				
0	their negative, but also of their positive liberty.				
Civil/Political liberties	Accountability				
Authorization	Legitimacy				
Responsiveness	Accountability				
Transparency	Accountability				
Direct Participation	Legitimacy				
Quality of Discourse	Legitimacy				
Representation	Legitimacy				
Effectiveness	Accountability				
Fairness	Accountability				
Vertical accountability	Accountability				
Horizontal Accountability	No need to be addressed in the current research, since the				
	MoUs do not, at least directly, affect the competition be-				
	tween different election candidates.				
Traditional Legitimacy	No need to be addressed in the current research, since the				
	Greek political system does not particularly belong to this				
	model				
Charismatic Legitimacy	No need to be addressed in the current research, since the				
	Greek political system does not particularly belong to this				
	model				
Legal-Ration Legitimacy	Legitimacy				
Adjudicative Legitimacy	No direct relation to this research				
Social Legitimacy	No direct relation to this research				
Normative Legitimacy	Belongs in a large part to Participatory Legitimacy				
Participatory Legitimacy	Legitimacy				
Democratic Legitimacy	Accountability				
Actual Consent	No direct relation to this research, since it usually implies				
	direct participation tools, such as referenda				
Output legitimacy	Accountability				
Input legitimacy	Legitimacy				
Subdivision of Legitimacy	Serves as a framework of the research - legitimacy to be				
according to the objects it	investigated according to political decisions and political				
investigates	authorities				

In the above table, all concepts utilized in this research and relevant to liberal democracies are narrowed down to legitimacy and accountability. From the literature review, it is evident that these two concepts are the most core and fundamental to liberal, Western democracies. In this research, accountability and legitimacy will be utilized in their general forms, so as to include all the more specific concepts in relation to liberal democracy. The purpose of this is that the constitution is meant to "establish the basic framework of validity of all other legal norms, giving concrete legal form to the basic values which underpin democracy as a political form, and to the key socio-economic choices made by the political community" (Menendez 2009, 285).

Based on the above, within this investigation, democratic value is operationalized as accountability and legitimacy. Accountability in this research is defined as provisions that the citizens can hold their rulers accountable for, most often on an ex post facto basis (elections). Legitimacy in this research is defined as provisions that are laid down for the government to follow necessarily and must not violate. Therefore, in regards to the MoUs, the criteria with which the affected policy areas are separated between legitimacy and accountability are evident. Policies that belong to the former are fundamentally within the authority of an elected government, both for the decision and the execution parts. Concepts that belong to the latter are in relation to what a government delivers to the people, which might be legitimate but just not good enough.

Due to the above, it cannot be expected that the constitution will necessarily refer to output legitimacy or fairness. However, by grouping all the specific individual concepts under either accountability or legitimacy, the investigation can effectively analyze the CG. Since the CG is meant to guarantee not the specifics, but rather the most fundamental and general democratic principles, accountability and legitimacy, the two core fundamental principles of liberal democracies, are the right concepts through which to filter the CG. Given the operationalization of the aforementioned two concepts throughout the Literature Review section, as well as the Research Methods section of this research, it is to be expected that all democratic principles within the CG can be filtered down to either accountability or legitimacy. The MoUs can then be measured against the same basic criteria, and thus establish whether they negatively affect the democratic value of Greece. The concepts of accountability and legitimacy are utilized so as to capture not so much the letter but the spirit of the law behind the CG, and investigate how that is affected by the MoUs.

The method utilized in this research is comparative, in depth analysis, of the MoUs in contrast to the CG. The first section of the analysis will separate all the principles of legitimacy and accountability within the CG. The second section will isolate all the parts of the MoUs that seem to violate either legitimacy and/or accountability. The comparison will be highlighted and conclusions will be reached regarding the conflict between the two sets of documents.

4 Analysis

To begin the analysis, all the different principles of legitimacy and accountability in the constitution are provided, in the following table:

Table 5: Legitimacy and Accountability Principles in the CG (Hellenic Republic 2008)

	Article	Section	Democratic Principle	Sub-category of Democratic Principle	
--	---------	---------	----------------------	--------------------------------------	--

1	1	Legitimacy	Form of Government.
1	2	Legitimacy	Electorate Independence.
	$\frac{2}{3}$	Accountability	People have the Power, which is to be exercised by
	-	,	the CG. Human Rights.
2	1	Accountability	
4	1	Accountability	Equality Gender Equality.
	2	Accountability	Proportional Participation to public charges.
-	5	Accountability	Processeria in the second for the second sec
5	1	Accountability	Free participation in social / political / economic life Participation not to infringe on the rights of others.
16		Legitimacy	Education.
16	2	Accountability	Education.
	4	Accountability	
17	1	Accountability	Protection of Private Property.
21	3	Accountability	Provision of Health care.
22	1	Accountability	Right of Employment.
	2	Legitimacy	Employment Provisions. Public Service Employment Provisions.
	3	Legitimacy	r uone Service Emproyment Provisions.
23	2	Accountability	Right to strike.
25	1	Legitimacy	Human Rights – Welfare State.
-	2	Accountability	Human Rights / Social Progress.
26	3	Legitimacy	Judiciary to exercise judicial power in the name of the
-			Greek People.
28	2	Legitimacy.	International Agreements Validation.
51	2	Legitimacy	MPs are the state's representatives. Election by direct, secret ballot by all citizens, over
	3	Legitimacy	the legal age.
52	Article	Legitimacy	State to guarantee free expression of public will.
	single	0 1	
62	Article	Legitimacy	MPs not prosecuted for crimes during their office
	Single		term.
73	1	Legitimacy	Parliament retains the right to initiate legislation.
	3	Legitimacy	Laws that result to a loss for the State / local govern- ment in order for payment to be made to a person
			cannot be discussed.
76	4	Legitimacy	Provisions for bills labeled as very urgent to be voted
	5	Legitimacy	by the Plenum. Government can request number of sittings that an
			urgent law will be debated.
78	2	Accountability	No financial charge can be imposed on a retroactive
0.2	1	-	effect. Government is responsible for policy with Greece, to
82	1	Legitimacy	be conducted according to the CG.
84	1	Legitimacy	Vote of confidence of the Parliament for the Govern-
	· ·	<i>c ,</i>	ment provisions.
	2	Legitimacy	Vote of no confidence of the Parliament for the Government provisions.
85			Ministers and Undersecretaries are responsible for
05	Article Single	Legitimacy	actions on behalf of the State, and no Presidential Act
6-	Ű	* •/•	can relieve them of liability for said actions.
87	1	Legitimacy	Judges to enjoy personal / functional independence when administering justice.
		A	Judges to have as their sole compass the laws of the
	2	Accountability	state and the CG, and may not violate them by a
00	1	Logitimoor	decision. Magistrates cannot partake in the Government
89	4	Legitimacy Legitimacy	Civil servants are the executive of the State and
103	1	Legitimacy	should serve the people and the CG
			Civil servants are permanent for their position unless

	4	Accountability	the position ceases to exist, and they may not be transferred without their opinion, lowered in rank or dismissed without a 2/3 decision of all permanent civil servants, against which decision they can appeal to the Supreme Administrative Court.
106	1	Accountability	State should place primary focus on economic devel- opment of all national economy sectors, especially
	2	Accountability	those that are in remote areas. Private initiative can not be allowed at the cost of human dignity or at the expense of the national economy.
120	2	Accountability	All Greeks need to demonstrate devotion to the CG and the laws that originate from that.
	4	Legitimacy	All Greeks have to observe the application of the CG and must resist its possible abolition by anyone.

To conduct the comparison between the CG and the MoUs, a summation of all the policies affected by the MoUs is presented in the Table below, categorizing said policies according to the aforementioned criteria into Legitimacy or Accountability. Given the multifaceted and extremely detailed nature of the MoUs, the policy areas covered had to be grouped under certain summative titles. In Appenix B, the specific policies and MoU references that correspond to the titles presented in the Table below, are provided. Additionally, in Appendix C, the specific page references for each MoU with the corresponding general, summative policy titles are also supplied.

(Table Legend: X the MoU involves that policy / - the MoU does not

Table 6: Policy areas & MoUs

Policy Areas	Accountability / Legitimacy	3/5/10	6/8/10	8/12/10	28/2/11	4/7/11	30/11/11	1/3/12	9/3/12	13/12/12
Budget / Expenditure	Legitimacy	X	Х	Х	Х	Х	Х	-	Х	Х
Taxes / Revenues	Legitimacy	X	Х	Х	Х	Х	Х	-	Х	Х
Public Remuneration	Accountability	X	Х	Х	Х	Х	Х	-	Х	Х
Public Administration / Employment	Accountability	X	Х	Х	Х	Х	Х	-	Х	Х
Defense	Legitimacy	-	-	Х	-	Х	-	-	Х	Х
Energy	Accountability	X	Х	Х	Х	Х	Х	-	Х	Х
Labor Market / Business Environ-	Accountability	-	Х	Х	Х	Х	Х	-	Х	Х
ment										
Healthcare & Pharmaceuticals	Accountability	X	Х	Х	Х	Х	Х	-	Х	Х
Education	Accountability	X	Х	Х	Х	Х	Х	-	Х	Х
Judiciary	Legitimacy	-	-	-	-	-	Х	I	Х	Х
Tourism	Accountability	-	-	Х	-	Х	Х	I	1	-
Transportations	Accountability	X	-	Х	-	Х	Х	-	Х	Х
Banks & Bank of Greece (BoG)	Legitimacy	X	-	Х	Х	Х	Х	1	Х	Х
R&D	Accountability	X	-	Х	Х	Х	Х	-	Х	-
Procurement	Legitimacy	X	Х	-	-	Х	Х	-	Х	Х
Privatization	Legitimacy	X	Х	Х	Х	Х	Х	-	Х	Х
Legislation amendments /adoptions	Legitimacy	X	Х	Х	Х	Х	Х	-	Х	Х
Social Welfare / Spending	Accountability	X	-	Х	Х	Х	Х	-	Х	Х

involve that policy)

By comparing Table 5 and Table 6, it is evident that both the accountability and legitimacy of the CG are directly affected by the MoUs. All the policies existent in the MoUs require alterations in either the legitimacy or the accountability of the Greek state. The effect becomes amplified by the necessity and urgency with which these loans are presented to the Greek citizens. Given that necessity, the Greek state seems to be in a very disadvantageous position to negotiate or bargain different terms of the MoUs. This lack of resistance, along with the aforementioned necessity and urgency, greately enhance the negative effect of the MoUs on the democratic principles of the CG. Finally, the number of policies affected by the MoUs (18 policy group titles which include numerous more specific policy areas within them and with some being considered core state policy - taxation, budget, public employment, justice, etc), as well as the level of detail with which they are shaped in the MoUs, are staggering. It almost seems as if the Greek government is left, not only with no choice but to adopt the MoUs (on account of the difficult financial situation as presented in the Greek people), but is also left with no policy to conduct, as all policy areas are covered at tremendous length and in extreme detail by the MoUs. The Greek MoUs present an excellent guideline to building an entire state from the ground up all over again, with absolutely no legitimacy or accountability whatsoever.

5 Conclusion

This paper investigated whether the Greek MoUs affect the democratic value of the Greek state, as enshrined in the CG, the most fundamental democratic document of any democratic country (of a civil law tradition). In this investigation, the two fundamental ideals of democracy in the Western, liberal democratic state were isolated in relevance to related literature, namely accountability and legitimacy. They were operationalized and tracked down within the CG. All the policy areas that are affected by the MoUs were presented in titled groups, and their effect on legitimacy or accountability was evaluated. It was found that the MoUs have a drastic effect on the democratic values of the Greek state, as enshrined in CG broken down into the two most very basic concepts of democracy, i.e. accountability and democracy. Not only that, but the great variety, detail and number of policies outlined by the MoUs, as well as the urgency and necessity with which they are presented, greatly enhance this democratic degrading.

Haggard (1985) relates structural adjustment programs to democracy, arguing that "the very nature of the democratic system and legal guarantees it affords limit the state's ability to act" (512). He further suggests that due to political competition within a modern democratic system, opposition parties might gain popularity by opposing structural adjustment measures resulting in the danger of destabilization. There is evidence that the mere effort to structurally adjust the economies of countries has been proven to be closely related to "installation of authoritarian governments" (Haggard 1985, 509-12). Remmer (1986) also goes on to argue that a democratic regime is a harder framework within which austerity programs can be applied, rather than military dictatorships for example. The above is due to the democratic state's transparent organization, strong bureaucracy, strong technocratic control, fragmented decisionmaking procedures that allow multiple political actors to interfere (3).

However, based on the current crisis, it could be argued that these disadvantages have been turned to advantages for implementation of structural adjustment programs within democracies. Transparent organization coupled with fragmented decision-making can yield an endless web of actors and offers great grounds for endless shifting of blame, technocratic control can become useful in drafting policy, while strong bureaucracy lends itself to immediate control of policy implementation. In addition, democratic regimes, when implementing said programs and with the assistance of the mass media, can appeal to the greater national good and create a positive reinforcement for the entire population in calls for unity against financial destruction (Remmer 1986, 3).

Therefore, from the above investigation, it can be concluded that the democracy of the Greek state, based on its two most fundamental principles, that is accountability and legitimacy, as presented in the most fundamental democratic document of the Greek state, that is the Constitution of Greece, has suffered a great loss of democratic value on account of the MoUs. The kind, number and intensity with which these policies are adopted, leads to a democratic deficit of Greece.

6 References

6.1 MoUs

European Commission (2010) "The Economic Adjustment Programme for Greece," http://ec.europa.eu/economy_finance/publications/occasional_paper/2010/pd f/ocp61_en.pdf (April 10, 2013)

European Commission (a) (2012) "The Second Economic Adjustment Programme for Greece: First Review – December

2012,http://ec.europa.eu/economy_finance/publications/occasional_paper /2012/pdf/ocp123_en.pdf, (April 10, 2013)

European Commission (b) (2012) "Memorandum of Understanding between the European Commission Acting on Behalf of the Euro Area Member States and the Hellenic Republic,"

http://ec.europa.eu/economy_finance/eu_borrower/mou/2012-03-01-greecemou_en.pdf, (August 10, 2012)

IMF (a) (2010) "Letter of Intent, Memorandum of Economic and Financial Policies, Technical Memorandum of Understanding, and Memorandum of Understanding on Specific Economic Policy Conditionality (European Commission and European Central Bank)," http://www.imf.org/external/np/loi/2010/grc/080610.pdf,, (August 10, 2012)

 IMF (b) (2010) "Letter of Intent, Memorandum of Economic and Financial Policies, and Technical Memorandum of Understanding,"

http://www.imf.org/external/np/loi/2010/grc/120810.pdf, (August 10, 2012)

IMF (c) (2011) "Letter of Intent, Memorandum of Economic and Financial Policies, and Technical Memorandum of Understanding," http://www.imf.org/external/np/loi/2011/grc/022811.pdf, (August 10, 2012) IMF (d) (2011) "Letter of Intent, Memorandum of Economic and Financial Policies, and Technical Memorandum of Understanding," http://www.imf.org/external/np/loi/2011/grc/070411.pdf. (August 10, 2012)

```
    IMF (e) (2011) "Letter of Intent, Memorandum of Economic and Financial Policies,
and Technical Memorandum of Understanding,"
http://www.imf.org/external/np/loi/2011/grc/113011.pdf, (August 10, 2012)
```

- IMF (f) (2012) "Letter of Intent, Memorandum of Economic and Financial Policies, and Technical Memorandum of Understanding," http://www.imf.org/external/np/loi/2012/grc/030912.pdf (August 10, 2012)
- IMF (g) (2012) "Letter of Intent, Memorandum of Economic and Financial Policies, and Technical Memorandum of Understanding," http://www.imf.org/external/np/loi/2012/grc/122112.pdf (January 6, 2013)

IMF (2013) "Letter of Intent, Memorandum of Economic and Financial Policies, and Technical Memorandum of Understanding," http://www.imf.org/external/np/loi/2013/grc/071713.pdf, (September 9, 2013)

6.2 Journals & Books & Laws

Beetham, David (1999) Democracy and Human Rights. Cambridge: Polity Press

- Beetham, David & Lord, Christopher (1998) *Legitimacy and the EU*. Essex: Addison Wesley Longman Limited.
- Bellamy, Richard (2010) "Democracy without democracy? Can the EU's Democratic 'Outputs' be Separated from the Democratic 'Inputs' provided by competitive parties and majority rule," *Journal of European Public Policy* 17(1): 2-19
- Birch, Anthony (2004) Ancient Greek Democracy, Routledge, London, UK
- Brassett, James & Tsingou, Eleni. 2011. "The Politics of Legitimate Global Governance," *Review of International Political Economy* 18(1): 1-16.
- Broome, Andre & Clegg, Liam & Rethel, Lena (2012) "Global Governance and thePolitics of Crisis," *Global Society* 26(1): 3-17.
- Cheneval, Francis & Schimmerlfennig, Frank (2013) "The Case for Demoicracy in the European Union," *Journal of Common Market Studies* 51(2): 334-350
- Craigwell, Roland & Lorde, Troy & Moore, Winston (2013) "Fiscal policy and the duration of financial crises," *Applied Economics* 45(6): 793-801
- Follesdal, Andreas (2006) "Survey Article: The Legitimacy Deficits of the European Union," *The Journal of Political Philosophy* 14(4): 441-468
- Guastaferro, Barbara & Moschella, Manuela. 2012. "The EU, the IMF, and the Representativ Turn: Addressing the Challenge of Legitimacy," Swiss Political Science Review 18(2): 199-219
- Hansen, Mogens Herman. 2007. "The Ancient Athenian and the Modern Liberal View of Liberty as a Democratic Ideal," in *Ancient Greek Democracy: Readings and Sources* eds Eric Robinson, Blackwell Publishing: Massachusetts, USA

Held, David (2006) Models of Democracy, Politiy Press, Cambridge, UK

Hellenic Republic (2008) The Constitution of the Greece. Athens

- Hobolt, Sara Binzer (2012) "Citizen Satisfaction with Democracy in the European Union," *Journal of Common Market Studies* 50(S1): 88-105.
- Huller, Thorsten (2012) "On Infeasibilities of Cosmopolitan Democracy Lesson from the European Union," *Swiss Political Science Review* 18(2): 249-271.
- Law of the Hellenic Republic(2011) "Law No 3986: Emergency Measures of the Implementation of Midterm Framework of Fiscal Strategy 2012-2015," *Official Journal of the Hellenic Republic*: 152 A/01-07-2011.
- Martins, Nuno (2012) "Globalisation, Inequality and the Economic Crisis," *New Political Economy* 16(1): 1-18
- Menendez, Agustin Jose (2009) "The European Democratic Challenge: The Forging of a Supranational Volonte Generale," European Law Journal 15(3): 277-308
- Neyer, Jurgen (2010) :Justice, Not Democracy: Legitimacy in the European Union," Journal of Common Market Studies 48(4): 903-921
- Olsen, Johan P (2013) "The Institutional Basis of Democratic Accountability," West European Politics
- Robinson, Eric. W (2007) Ancient Greek Democracy: Readings and Sources, Blakcwell Publishing, Massachusetts, USA
- Saward, Michael (1994) "Democratic Thoery and Indices of Democratization," In *Defining and Measuring Democracy*, eds. David Beetham, London: SAGE Publications Ltd.
- Schmidt, Vivien A (2013) "Democracy and Legitimacy in the European Union Revisited: Input, Output and Throughput," *Political Studies* 61: 2-22
- Sinclair, Timothy J (2010) "Round Up the Usual Suspects: Blame and the Subprime Crisis," *New Political Economy* 15(1): 97-107
- Wallison, Peter J (2010) "The Causes of the Financial Crisis and its Consequences," Presented at the Independent Petroleum Association of America, Colorado Springs, USA.
- Weir, Stuart & Beetham, David (1999) Political Power and Democratic Control in Britain. London: Routledge
- Wimmel, Andreas (2009) "Theorizing the Democratic Legitimacy of European Governance: A Labyrinth with No Exit?" *Journal of European Integration* 31(2): 181-199.

6.3 Online sources

- Al Jazeera (2012) "Cyrpus Asks EU for Financial Bailout: The Fifth Eurozone Nation to Seek Help Says It Wants to Limit Expoure of I Financial Institutions to Greece Crisis," *AlJazeera* http://www.aljazeera.com/news/europe/2012/06/ 201262517189248721.html (August 11, 2013)
- Council of State (a) (2012) "The Council of State," http://www.ste.gr/FL/main_en.htm (August 10, 2012) Council of State (a). 2012.
- Council of the State (b) (2012) "CoS Plenary 668/2012," http://www.dsanet.gr/Epikairothta/Nomologia/668.htm (April 3, 2013)
- Greek Ministry of the Interior (2009) "National Elections 2009," http://ekloges-prev.singularlogic.eu/v2009/pages/index.html?lang=en (April 1, 2013)

- Investingbonds.com (2010) "Types of Bonds: What are Asset-Backed Securities?" http://www.investinginbonds.com/learnmore.asp?catid=5&subcatid=16&id= 10 (April 10, 2013)
- Investopedia (2012) "Definition of Collateralized Debt Obligation CDO," http://www.investopedia.com/terms/c/cdo.asp (October 10, 2012)
- Investopedia (2013) "Definition of Alt-A," http://www.investopedia.com/terms/a/alta.asp (August 10, 2013)
- Kollewe, Julia (a) (2010) "Ireland's Debt Downgraded by Credit Ratings Agency," *theguardian*, July 19, http://www.guardian.ac.uk/business/2010/iul/10/ireland.debt.downgraded

http://www.guardian.co.uk/business/2010/jul/19/ireland-debt-downgraded-rating-agency (April 2, 2013)

- Kollewe, Julia (b) (2010)"Euro Continues to Fall as EU Fails to Agree Greece Rescue Plan," *theguardian*, February 12, http://www.guardian.co.uk/business/2010/feb/12/euro-continues-to-fall (April 2, 2013)
- Mamundi, Sam (2008) "Lehamn Folds with Record \$613 Billion Debt," *MarketWatch*, September 15, http://articles.marketwatch.com/2008-09-15/news/30748651_1_lehman-bonds-debt-lehman-brothers-holdings (April 1, 2013)
- Mead, Nick & Blight, Garry. 2013. "Eurozone Crisis: Three-And-A-Half Years of Pain," *theguardian*, March 28, http://www.guardian.co.uk/business/interactive/2012/

oct/17/eurozone-crisis-interactive-timeline-three-years (April 1, 2013)

- Moya, Elena (a) (2010)"Greece Asks for IMF-EU Rescue Talks," *theguardian*, April 15, http://www.guardian.co.uk/business/2010/apr/15/greece-imf-rescue1 (April 2, 2013)
- Moya, Elena (b) (2010)"Spain Seek to Calm Fears It Is 'Next Greece' as European Market Plunge," *theguardian*, February 5, http://www.guardian.co.uk/business/2010/feb/05/spain-bid-calm-turmoil (April 2, 2013)
- Ryssdal, Kai (2012) "Greek Bond Investors Take Big Haircut in Bailout Deal," *Market Place* http://www.marketplace.org/topics/world/greek-debt-crisis/greek-bond-investors-take-big-haircut-bailout-deal (August 10, 2013)
- theguardian.co.uk (2010) "Greek Protesters Riot in Athens," *theguardian*, May 1, http://www.guardian.co.uk/world/gallery/2010/may/01/greek-protesters-riotathens (April 2, 2013)
- The Tribune (a) (2013) "Fitch on Cyprus: Easier to Negotiate with Anastasiadis," *The Tribune*

February 26. http://www.tovima.gr/politics/article/?aid=500110 (August 9, 2013)

- The Tribune (b) (2013) "W. Schäuble: Cyprus, Brussels and the ECB are responsible for the haircut of deposits under €100.000," The Tribune March 18. http://www.tovima.gr/finance/article/?aid=503503 (August 9, 2013)
- The Tribune (c) (2013) "Schäuble Ultimatum to Cyprus," The Tribune March 24. http://www.tovima.gr/finance/article/?aid=504311 (August 10, 2013)
- topetopetopapagalos (2010) "Λεφτά Υπάρχουν Χρήματα Υπήρχαν," http://www.youtube.com/watch?v=WmPm1vb8q1c (April 2, 2013)

- U.S. SEC Securities and Exchange Commission (2010) "Mortgage-Backed Securities," http://www.sec.gov/answers/mortgagesecurities.htm (October 10, 2012)
- U.S. SEC Securities and Exchange Commission (1994) "Nationally Recognized Statistical Rating Organizations," http://www.sec.gov/rules/concept/34-34616.pdf (October 15, 2012)
- Xarhs123 (2012) "Cypriot Banks in the Aftermath of the Greek Haircut," *the Cyprus Lawyer: The Life and Practice of a Lawyer in Cyprus*, http://thecypruslawyer.com/cypriot-banksin-the-aftermath-of-the-greek-haircut/cyprus-news (August 3, 2013)

7 Appendix

7.1 Appendix A: Table 7: €MS /EUMS under EU/IMF Financial Assistance and their respective

MoUs, in order of declining amount of financial assistance provided

(European Commission 2013 & also IMF country homepage, for each of the specific countries)

Country	Total Assistance	Date of MoU	No of	Organiza- tion	Total Pages
		02.05.2010	Pages		
		03-05-2010	92	EU	
		06-08-2010	54	IMF	
		08-12-2010	62	IMF	
_		28-02-2011	59	IMF	
Greece	€ 140.8 bln	04-07-2011	89	IMF	1021
		30-11-2011	77	IMF	
		01-03-2012	12	EU	
		09-03-2012	100	IMF	
		13-12-2012	282	EU	
		21-12-2012 ⁸	145	IMF	
		17-07-2013 ⁸	49	IMF	
		03-12-2010	37	IMF	
		02-2011	89	EU	
		10-02-2012	36	IMF	
Ireland	€ 85 bln	12-03-2012	36	IMF	341
		30-05-2012	33	IMF	
		20-08-2012	33	IMF	
		29-11-2012	43	IMF	
		03-06-2013	34	IMF	
		17-05-2011	68	IMF	

⁸ Not covered within this research

			06-2011	1	15	F	U		
			01-09-2011		6		0 1F		
Portugal	€ 78 bln		09-12-2011	75 61			1F 1F		607
i oi tugai	C /0 biii		27-06-2012				MF		007
			14-10-2012	54			1F 1F		
			19-12-2012				1F 1F		
			15-03-2012 12-06-2013	66 IN 24 IN		1F 1F			
				1					
Spain	€ 41.3 bln		10-2012	6	9	E	U		69
			24-04-2009			22	IM		
			23-06-2009	15			EU		_
			08-09-2009			21	IM		{
			05-02-2010		19 8		IM		1
		25-02-2010 16-06-2010 20-07-2010 09-09-2010		24			EU		1
							IM EI		1
				7 19		EU IMF			
						19			
		22-12-2010 19-01-2011 10-03-2011					IM EU		-
Romania					28	IM		422	
Komama	€ 20 bln		08-04-2011			7	EU		422
		09-06-2011				29	IMF		1
		14-09-2011				29	IM		
			02-12-2011			35	IM		
			27-12-2011			9	EU		
			28-02-2012			39	IM		
			08-06-2012			38	IM	F	
			29-06-2012			9	EU	J]
			12-09-2012			35	IM	F	
			10-06-2013			40	IM	F	
			04-11-2008			13	IM	F	
Hungary	€ 14.2 bln		16-09-2009			18	IM	F	65
			04-12-2009			17	IM	F	
			04-03-2010			17	IM	F	
Cyprus	€ 10 bln		29-04-2013			64	IM	F	221
			16-05-2013			57	EU	J	
			28-01-2009			12	EU	J	
			13-07-2009			11	EU	J	
			22-01-2010			24	IM	F	
			22-02-2010			11	EU	J	
Latvia	€ 7.5 bln		05-07-2010			22	IM	F	159

20-07-2010	12	EU	
09-05-2011	37	IMF	
07-06-2011	14	EU	
21-12-2011	16	EU	

7.2 Appendix B

Table 8: General Policy Titles and Specific policy areas covered

Policy Category	Specific policies under that category
Banks / Bank of Greece	Bank loans
(BoG)	Bank mergers/acquisitions
(In Tables 9i & 9ii appear-	Banks' resolutions
ing as: Banks)	 BoG reforms (payroll, pensions, structure, role, etc)
Budget / Expenditure	 Direct changes / alterations in the state, local government, or state enterprises budget
(In Tables 9i & 9ii appear-	
ing as: Budget)	• Direct guidelines for the implementation of all the above budgets.
ing us. Budget)	• Extra-budgetary funds' expenditure
DC	Government expenditure / arrears
Defense	• Defense / military spending (ex. equipment)
(In Tables 9i & 9ii appear-	
ing as: Defense)	
Education	Franchised diplomas
(In Tables 9i & 9ii appear-	Professional qualifications (recognition, validation, etc)
ing as: Education)	
Energy (In Tables Qi & Qii enneer	• Electricity and gas policies, tariffs, etc
(In Tables 9i & 9ii appear- ing as: Energy)	• Electricity and gas agencies' changes
ing as. Energy)	Liberalize / open the energy market to third parties
	Policies pertaining to gas and electricity companies
	Renewable energy
Healthcare & Pharmaceuti-	Health funds
cals	 Hospital Restructuring (accounting, pricing, etc)
(In Tables 9i & 9ii appear-	 Management of pharmaceuticals (ex. generics / off-patent over brand)
ing as: Healthcare)	 Management of hospitals
	 Alterations in hospital staff regime
	Hospital computerization
	 Hospital accounting practices reform
	Pharmacies' policies
Judiciary	 Tax administration disputes system alterations
(In Tables 9i & 9ii appear-	• Trial fees
ing as: Judiciary)	
Labor Market	 Liberalize / open closed and restricted professions
&	Reorganize / alter the private sector workplace
Business Environment	 minimum wage / wage agreements / wage controls
(In Tables 9i & 9ii appear-	 collective agreement regime,
ing as: Labor)	o overtime,
	 employment protection laws,
	 part-time vs full-time contracts,
	 sectoral agreements seniority
	 list of heavy and arduous professions
	 competitiveness issues, etc
	Alterations in the business environment (setting-up / resolution of companies, enterprise
	licensing, market controls, e-platforms, etc)
	Policies on competition
	• FDI
	· Public works / investments (decrease delays, simplification of processes, etc)
	Circulation fees for cars
	Point of Single Contact
	• Changes to all private sectors (retail, communications, etc), save for energy, tourism
1	and transportation, that are presented separately below.
	and autoportation, that are presented separately below.

	D1 1 1D1 1 1
	• Digital Dividend
	• Unemployment (benefits, time schedules)
9	Imports / Exports policies
Legislation ⁹	Adopt new legislation (laws and statutes)
(In Tables 9i & 9ii appear-	Repeal / alter old legislation
ing as: Legislation)	Issue new statutory regulations (Presidential Decrees, Ministerial Decisions, Circulars)
	 Creation of new agencies, offices, Secretary General positions, registries etc
Privatization	• Privatization through concessions, sales, tenders of anything public (real estate, airports,
(In Tables 9i & 9ii appear-	ports, gaming, buildings, banks, utilities, etc)
ing as: Privatization)	 Allocation of profits from privatizing assets
Procurement	Any type of public procurement, for with pharmaceuticals, hospitals, public works, e-
(In Tables 9i & 9ii appear-	auctions, etc
ing as: Procurement)	
Public administration &	Public employment policy
Employment	 staff reductions,
(In Tables 9i & 9ii appear-	◦ labor reserve,
ing as: Administration)	 mobility scheme,
	 functional reviews, etc
	• Revenue administration reform (closure of offices, change of type of agency - state vs
	independent – staffing plans, object and number of audits)
	State-Owned Enterprises & Public entities policies (restructuring, mergers, price
	controls, closures, etc)
	Statistical overhaul (rules and obligations, agencies' restructuring, etc)
	Local government restructuring / reform
Public remuneration	Pension reform / cuts (including pension funds)
(In Tables 9i & 9ii appear-	 Public wage bill / cuts (including special regimes' wages)
ing as: Remuneration)	 Fublic wage off / cuts (including special regimes wages) SOEs wage changes
ing us. Remaneration)	
Conicl Connite & Welford	Remuneration system changes
Social Security & Welfare	Social benefits (ex. family allowances)
(In Tables 9i & 9ii appear-	Nets of protection for the most vulnerable
ing as: Welfare)	Social programs
	 Social Security Funds & their expenses
	Unemployment benefits
	Social contributions' regime
	List of disabilities
Taxes / Revenue	• VAT
(In Tables 9i & 9ii appear-	Wage taxation
ing as: Tax)	Regular / excise taxes on fuel, tobacco and alcohol
	Luxury taxes
	• Green tax
	Presumptive taxation
	• Levies
	Gaming licenses
	Environmental licensing
	Telecommunication licenses
	Concession licenses
	• Real estate taxes
	• Unauthorized construction fines
	Incentives to regularize the violations of use of land
	Customs changes
	 Tax collection systems (penalties, fines, etc – except for judicial dispute resolution,
	which is assigned under the title 'Judiciary' presented above)
	Solidarity tax
Tourism	Tourism sector changes
(In Tables 9i & 9ii appear-	Tourist coaches changes
ing as: Tourism)	
Transportation	Road freight transport / road haulage
(In Tables 9i & 9ii appear-	Railway sector reforms / alterations / liberalization
ing as: Transportation)	Buses, coaches, limousines, trucking industry, taxis, ferries, ports
	· · · · · · · · · · · · · · · · · · ·

⁹ Using the distinction of Menendez (2009) between the constitution, statutes, and statutory regulations (284-6)

7.3 Appendix C

Table 9i: Policy categories and page references¹⁰

Pag- es	Policies	Pag- es	Policies	Pages	Policies	Pages	Policies
MoU 3	3/5/10 (Europe- nmission 2010)	MoU	6/8/10 [IMF) 2010]	MoU 8/12 (b) 2	2/10 [IMF		2/11 [IMF (c) 011]
41-2	Welfare Administra- tion Remunera- tion	1-2	Labor Remunera- tion	3	Transpo rt	4	Asministra- tion
42	Welfare	3	Expendi- ture Remunera- tion	8	Tax Labor	5, 31, 51	Tax Remunera- tion Administra- tion
43	Administra- tion Tax	4	Budget	9,10, 52	Healthc are Remu- neration Welfare Tax	6, 43	Welfare Labor
44	Legislation Remunera- tion Tax	5	Remunera- tion Healthcare Admin- istration	10	Tax Admin- istration Remu- neration Defense	7, 31	Healthcare Tax Judiciary Legislation
45	Legislation	6	Pension policy Legisla- tion	11	Tax Judicial Admin- istration	9, 32, 51	Privatization
46	Banks	7	Banks Legisla- tion Labor	12	Admin- istration Remu- neration Banks	10, 11	Banks
47	Banks	8, 38, 46, 48	Privatiza- tion Legisla- tion Labor	13,14	Banks Legisla- tion	12, 47, 52	Labor Tourism
48	Remunera- tion Healthcare Administra- tion Legislation	11	Admin- istration Remunera- tion Tax	14	Labor	15, 31	Remunera- tion Legislation Privatization Labor Banks Administra- tion
49	Labor	31	Budget	15, 40, 49	Labor Privati- zation	26, 55	Pension
50	Banks	32	Tax	18	Remu- neration Admin-	30	Budget

¹⁰ 1/3/12 MoU not in the tables since it only includes a commitment to the next MoU and EFSF loan financial requirement agreements

							istration		
							Tax		
							Legisla-		
							tion		
							Privati-		
							zation		
							Labor		
							Banks		
							Defense		
51	Tax		34	Labor		30-3,	Remu-	31	Tax
				policy		56	neration		
							Admin-		
							istration		
							Healthc		
							are		
							Welfare		
							Defense		
							Tax		
55	Tax	-	35	Admin-	-	33	Admin-	32, 43, 47,	Remunera-
55	1 0 1		55	istration		55	istration	52, 45, 47, 51	tion
				istration			Welfare	51	Administra-
							Admin-		tion
									1011
							istration Privati-		
	Deciliari		21	A due!		24	zation	22	A Justa 1 4
57	Budget		36	Admin-		34	Defense	33	Administra-
	Remunera-			istration			Budget		tion
	tion						Tax		Remunera-
	Tax						Legisla-		tion
	Legislation						tion		Transporta-
	Privatization								tion
	Administra-								
	tion								
59-	Budget		37	Budget		35	Healthc	34, 44-6,	Administra-
60	Tax						are	51, 55	tion
	Legislation								Healthcare
60	Banks		38	Budget		36	Banks	37, 47, 51	Labor
	Legislation						Legisla-		Legislation
	Remunera-						tion		0
	tion						Labor		
	Administra-								
	tion								
	Labor								
61	Labor		39	Remunera-		37, 53-	Labor	38, 47	Tourism
01	Legislation		57	tion		4, 57	Legisla-		Education
	Administra-			Budget		., .,	tion		
	tion			Duager			Admin-		
							istration		
62-3	Budget		40	Healthcare		38	Admin-	39, 48	Legislation
02-5	Tax		40	Tranneare		50	istration	37,40	Labor
	Legisla-						Legisla-		R&D
	tive reform						tion		
							Procure		
							cure-		
	-					a a	ment		
63-5	Tax		41	Labor		39, 56	Legisla-	41	Energy
	Remunera-						tion		Privatization
	tion						Tour-		
	Admin-						ism		
	istration								
65	Administra-		42,	Energy		40, 57	Privati-	42, 48, 52,	Education
	tion		44				zation	54	
66-7	Legislation		45	Labor		41, 49,	Tourist	43	Budget
	Banks					57	policy		
	Budget						Energy		
							policy		
							Labor		
	•								

						Energy			
67	Remunera- tion		46	Remunera- tion	42	R&D		44	Labor Banks
	Procurement Healthcare			Welfare Admin-					Defense Privatization
68	Labor	-	47	istration Admin-	43	Labor	-	50	Tax
	Legislation Administra- tion		.,	istration					Administra- tion Welfare
69	Education Labor Budget R&D		48	Tax	44	Tax Admin- istration		52	Labor
70	Energy Legislation Administra- tion Welfare		52	Legisla- tion	45-7, 52-3, 56, 57	Remu- neration Healthc are		53	Privatization
71	Labor Budget Administra- tion Tax Remunera- tion				47	Admin- istration			
72	Legislation				48, 53- 4, 57	Banks Legisla- tion			
75- 77	Legislation				49	Labor			
78-9	Procurement Remunera- tion Administra- tion				50, 54, 56	Energy Trans- porta- tion Educa- tion			
79- 81	Remunera- tion Administra- tion				51	Remu- neration			
81-2	Healthcare				55	Admin- istration Remu- neration Tax			
82	Transporta- tion Labor				56	Tax Labor			
83-4	Energy								
84	R&D								
90	Remunera- tion								

Table 9ii: Policy categories and page references

Pages	Policies	Pages	Policies	Pages	Policies	Pages	Policies
	7/11 [IMF 2011]	MoU 30/1 (e) 2	1/11 [IMF 011]		/ 12 [IMF (f) 012]		/12 [European on (a) 2012]
4	Admin- istration	5-6	Remu- neration Admin-	7-8	Remunera- tion Admin-	194-5	Budget Legislation Tax

	1		:			:-44:		Duinutination
			istration Tax			istration Healthcare		Privatization
			Tax			Welfare		
			D.C.	-	0		105	T.
5	Admin-	6-7	Defense		8	Budget	195	Tax
	istration		Budget			Admin-		
	Remu-		Remu-			istration		
	neration		neration					
	Welfare		Healthca					
			re					
			Welfare					
			Tax					
6	Labor	7-8, 46	Admin-		9	Defense	196-9	Tax
	Remu-		istration			Tax		Administra-
	neration		Remu-					tion
	Healthca		neration					
	re		Labor					
	Tax		Welfare					
			Healthca					
			re					
			Tax					
			Legisla-					
			tion					
7-8	Tax	8-9	Tax		10	Legislation	200	Budget
/-0	Admin-	0-7	Admin-		10	Banks	200	Administra-
	istration		istration			Admin-		tion
	istration		Legisla-			istration		Healthcare
	1		tion			istration		ricanneare
0	Adusta	0.45		-	11	A durin	201.2	Т
9	Admin-	9, 45	Legisla-		11	Admin-	201-2	Tax
	istration		tion			istration		Budget
			Tax			Tax		Administra-
			Admin-			Legislation		tion
			istration					Remuneration
13-4	Banks	10, 46	Legisla-		12	Welfare	202	Legislation
	Privati-		tion			Legislation		Budget
	zation		Admin-			Budget		Healthcare
			istration					
15	Privati-	15	Privati-		13	Admin-	203-4	Remuneration
	zation		zation			istration		Legislation
	Legisla-					Budget		Administra-
	tion							tion
								Banks
17-19,	Labor	16	Legisla-		14	Banks	204	Privatization
62-4,	Admin-		tion					Legislation
69-70	istration		Labor					Defense
	Transpor							
	t		1					
	Energy		1					
24	Admin-	17	Labor		15	Banks	205	Administra-
27	istration	1/	Tax		15	Legislation	205	tion
	Legisla-		101			Legislation		
	tion							Legislation
	Tax		1					
1			1					
	Defense							
	Legisla-		1					
1	tion		1					
1	Banks		1					
	Budget		1					
	Remu-		1					
	neration		L		L			L
25	Admin-	18-19,	Labor		16	Legislation	206-7	Administra-
	istration	61-2	Transpor			Banks		tion
	Labor		t					Legislation
	Tax		Energy					
	1		Judici-					
	ļ		ary					1
27	Admin-	19-20	Legisla-		19	Banks	207-8	Procurement

								T 1 .:
	istration			tion		Privatiza-		Legislation
	Privati-			Budget		tion		
	zation	-						
38-9	Tax		22, 42-3	Remu-	20-1	Privatiza-	208	Procurement
	Remu-			nerations		tion		
	neration			Admin-				
				istration				
				Legisla-				
				tion				
				Privati-				
				zation				
				Labor				
				Banks				
		-		Tax				
40	Remu-		23, 42,	Tax	21-2	Labor	209	Administra-
	neration		49	Legisla-		Legislation		tion
	Admin-			tion				Remuneration
	istration			Admin-				Welfares r
	Tax			istration				
				Welfare				
				Labor				
				Healthca				
	D 1	-	2.4	re		x 1	210	XX 1.1
44	Budget		24	Privati-	22	Labor	210	Healthcare
				zation		Legislation		Legislation
						Welfare		
46	Privati-		32	Budget	23	Labor	211-12	Healthcare
	zation			U U				
	Legisla-							
	tion							
47	Admin-	-	42	Tax	24	Labor	212-3	Healthcare
47	istration		42	Remu-	24		212-3	Tieanneare
	istration					Judiciary		
		-		neration				
48	Tax		43	Budget	25	Judiciary	213-4	Healthcare
	Judici-					Legislation		
	ary							
49-50	Admin-		44	Privati-	26	Budget	214	Healthcare
	istration			zation		e		
	Remu-			Banks				
	neration			Dunks				
	Pro-							
51.0	curement	-	4.5	x · 1	20	T.	01.5	XX 1.1
51-2	Labor		45	Legisla-	28	Tax	215	Healthcare
	Remu-			tion		Legislation		
	neration			Privati-				
				zation				
				Transpor				
1				t				
1				Tax				
52-58	Healthca		47	Admin-	29	Budget	216	Education
52-50	re		• /	istration	27	Remunera-	210	Legislation
								Legislation
1				Re-		tion		
1				munera-		Welfare		
1				tion		Education		
1						Admin-		
						istration		
						Healthcare		
						Legislation		
60	Banks		46	Judici-	30	Tax	217-8	Banks
				ary	-	Banks		
				Admin-		Legislation		
				istration		Judiciary		
61	Labor		48	Pro-	33-4	Legislation	218	Donka
01			48		33-4		218	Banks
	Admin-			curement		Labor		
	istration			Legisla-		Welfare		
				Legisla- tion		Welfare Privatiza- tion		

r			1					
						Banks		
						Judiciary		
						Tax		
						Admin-		
						istration		
						Remunera-		
						tion		
62	Labor	49	Labor	1	52	Budget	219	Banks
-			Remu-		-	Privatiza-		Legislation
			neration			tion		
64-5	Educa-	50-1	Bank		52-3	Healthcare	220	Banks
04-3	tion	30-1	Healthca		32-3		220	Daliks
						Defense		
	Labor		re			Budget		
						Remunera-		
						tion		
						Welfare		
66-8	Labor	51-2	Legisla-		53	Legislation	221	Banks
	Legisla-		tion			Remunera-		
	tion		Healthca			tion		
	tion		re			Budget		
			IC					
			_	_		Defense		
68	Tourism	53-5	Remu-		54	Privatiza-	222	Labor
1	Privati-		neration			tion		Legislation
	zation		I			Legislation		
70-1	R&D	56	Banks		55	Privatiza-	223-4	Labor
			1			tion		1
						Legis-		
						lation		
			_			Tax		
71-3	EU	57	Banks		56	Legislation	224	Legislation
	cohesion					Tax		Labor
	funds					Admin-		
	policy					istration		
73-5	Educa-	58	Labor		57	Tax	225-6	Labor
15 5	tion	50	Lucoi		57	Admin-	220 0	Legislation
	Judici-					istration		Budget
	ary					Judiciary		
77-80	Remu-	59	Labor		58	Budget	225-6	Legislation
	neration		Legisla-			Admin-		Labor
	Budget		tion			istration		
	Admin-		Tax			Legislation		
	istration					Welfare		
	Healthca					enure		
	re							
	Defense		1					1
	Welfare		1					1
	Tax							
V/////////////////////////////////////	X/////////////////////////////////////	60	Legisla-		59	Admin-	226-7	Legislation
	X/////////////////////////////////////		tion			istration		Tax
V/////////////////////////////////////	X/////////////////////////////////////		Labor					Administra-
V/////////////////////////////////////	X/////////////////////////////////////		Admin-					tion
<i>\////////////////////////////////////</i>	X/////////////////////////////////////							1011
	X/////////////////////////////////////	(2.2	istration		(2)	D	227.0	T 1
	X/////////////////////////////////////	62-3	Educa-		60	Procure-	227-8	Labor
V/////////////////////////////////////	X/////////////////////////////////////		tion			ment		Legislation
<i>\////////////////////////////////////</i>	X/////////////////////////////////////		Labor			Legislation		Tax
V/////////////////////////////////////	X/////////////////////////////////////		Tourism					1
V/////////////////////////////////////	X/////////////////////////////////////		Legisla-					1
V/////////////////////////////////////	X/////////////////////////////////////		tion					1
	X/////////////////////////////////////	64			61	Droguro	228.0	Labor
V/////////////////////////////////////	X/////////////////////////////////////	64	Admin-		61	Procure-	228-9	Labor
VIIIIIIIII			istration			ment		Legislative
V/////////////////////////////////////	X/////////////////////////////////////		Tourism			Legislation		Judiciary
						1 4 1 1		
			Transpor			Admin-		
						Admin- istration		
			Transpor			istration		
			Transpor			istration Remunera-		
		65	Transpor t	_	62.2	istration Remunera- tion	220 220	ludiciony
		65	Transpor		62-3	istration Remunera-	229-230	Judiciary Legislation

						Banks			
						Healthcare			
						Legislation			
						Welfare			
		66	R&D	-	63-4	Healthcare		230	Legislation
		00	Legisla-		05 1	ricultileure		250	Judiciary
			tion						succentry
		69-70	Educa-	-	65-6	Healthcare		231	Legislation
		09-70	tion		03-0	Tieatuicate		231	
									Energy
			Judici-						
		71.0	ary	-		XX 1.1	-		x
		71-2	Judici-		66	Healthcare		232	Legislation
			ary			Legislation			Energy
			Labor			Admin-			
				-		istration			
					67	Healthcare		233	Tax
						Admin-			Labor
						istration			Legislation
					68	Healthcare		234-5	Communica-
			V/////////////////////////////////////						tions policy
			V/////////////////////////////////////						Legislation
			V/////////////////////////////////////						adoption
					69	Procure-		235	Transporta-
						ment			tion
VIIIIII						Healthcare			Legislation
					70-1	Banks		236	Transporta-
					701	Dunks		250	tion
									Privatization
									Legislation
				-	71-2	Banks		237	Labor
					/1-2			237	Labor
				-		Legislation	-	220.0	x
					73	Legislation		238-9	Legislation
									Remuneration
				-					Labor
					74	Banks		239-40	Legislative
									Labor
									Education
					75	Labor		243	Legislation
						Legislation			Budget
					76	Labor		250-1	Remuneration
						Budget			Administra-
						Welfare			tion
						Tax			Healthcare
					77-8	Legislation		251-2	Administra-
						Labor			tion
									Budget
									Defense
			V/////////////////////////////////////						Healthcare
			X/////////////////////////////////////						Welfare
				-	78-9	Labor		252	Budget
			X/////////////////////////////////////		/0-9			232	Tax
			X/////////////////////////////////////			Legislation			Legislation
					70	Education			
			V/////////////////////////////////////		79	Education			X/////////////////////////////////////
					00.1	Legislation			X/////////////////////////////////////
			X/////////////////////////////////////		80-1	Labor			X/////////////////////////////////////
			V/////////////////////////////////////			Legislation			<u> </u>
			V/////////////////////////////////////		81-2	Legislation			X/////////////////////////////////////
						Labor			<u>X////////////////////////////////////</u>
			V/////////////////////////////////////		82	Transport			X/////////////////////////////////////
						Legislation			X/////////////////////////////////////
					83-4	Labor			X/////////////////////////////////////
						Legislation			X/////////////////////////////////////
			V/////////////////////////////////////			Tax			X/////////////////////////////////////
V/////////////////////////////////////						Transport			X/////////////////////////////////////
			v/////////////////////////////////////					<u>,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,</u>	<u>x////////////////////////////////////</u>
					84-5	Taxis			X/////////////////////////////////////
					84-5	Taxis Transpor-			

		tation	
		Privatiza-	
		tion	
		Energy	
	85	Privatiza-	
		tion	
		Energy	
	86-7	Energy	
		Legislation	
		Labor	
	87-8	R&D	
		Legislation	
	90	Education	
		Legislation	
		Judiciary	
		Labor	
	91	Judiciary	
<u> </u>		Legislation	

Children's Healthcare Facilities: Architectural aspects for the improvement of the environment's quality

Artemis Kyrkou¹, Prof. Fani Vavili²

¹ architect, PhD candidate Department of Architectural Design and Architectural Technology, School of Architecture, Aristotle University of Thessaloniki
 ² Dr architect, Professor at Department of Architectural Design and Architectural Technology, School of Architecture, Aristotle University of Thessaloniki

 ¹ amykyrkou@yahoo.gr, ²faniva@arch.auth.gr

The hospitalization process is inevitably a source of stress for an individual due to the unfamiliar setting and to the fear of pain and unknown medical test. When the patients are children it is even more crucial to offer a more secure and friendly hospital environment in order to eliminate as much as possible the stress factor. The design of healthcare facilities for children is a field of architecture that deeply affects the young mind and body of a child. During a hospital visit or stay the children's desires need to be fulfilled in order to promote healing, dismiss their negative feelings and restore balance. A space's quality is defined by its effectiveness, its efficiency and the satisfaction that can offer to its users. Basic parameters of architectural design such as the use of lighting and materials (both interior and exterior), the stability of an exterior design, the connection of a building to its surrounding and nature, orientation and wavfinding, scale and etc. can create a child friendly environment that will eliminate stress and promote healing. This dissertation is about the design of children's healthcare facilities and the way that architectural elements can create a more secure and less traumatic experience to hospitalized children. The purpose of this research is to specify the current trends of architecture in this field estimate the quality of children's hospitals and children's wards in the public general hospitals of Greece and propose a new design that focuses on a health promoting environment on the case study of an adolescent's psychiatry unit.

Keywords: children's healthcare, children's healing environments, children's psychology, environmental psychology

1 Introduction

Since the Modern Movement of architecture, hospital design has evolved and adjusted to the changing needs. Nowadays, it is considered to be a function of multiple variables. It is an architect's duty to design an environment (the term environment in this case includes all of its interpretations; natural, built, interior and exterior) that will propose security and will positively affect the patient. The hospitalization process is inevitably a source of stress for an individual. This is caused due to the absence of a familiar setting, the insecurity of the future, the fear of unknown medical tests or surgery, the pain, the restriction of social and everyday life that automatically affect the patient (Vavili, 2009). At an obvious level it is widely known that stress inhibits healing. Health is an overall condition that apart from the physical state depends on environmental, psychological, social and emotional factors. Therefore, the environment can have positive effects both on a visitor but more importantly on a recovering patient. The quality of healthcare facilities' environment is rather crucial for adults but it is even more significant when the patients are children. As Coull A. points out "Kids are vulnerable due to their developmental age. Hospitalisation is a defining experience for a child". In this case a major difference between adults and children is the way that they perceive the surrounding environment, due to the fact that they are still in a developmental stage (physically and spiritually).

2 Literature review

2.1 The role of the environment in children's hospitals

The design of healthcare facilities for children is a field of architecture that deeply affects the young mind and body of a child during its visit or stay. Therefore the elements of architecture that compose a facility's space (either public e.g. lobby, reception etc. or private e.g. patient's room, surgery etc.) should be carefully chosen and aim for the elimination of the stress and the feeling of security. "The architectural environment must not be an additional source of stress or add to the feeling of being out of control" (Verderber, 2010). Basic parameters of the design such as: the exterior of the facility (massing of horizontal or vertical elements, entrance), the interior (public spaces, corridors, reception etc.), the relationship between indoor and outdoor space, the lighting (natural or artificial), the materials (textures, forms, colors), the scale, orientation and wayfinding and acoustics determine the quality of the space. B. Lawson points out about the hospital design that "...architects should perhaps concentrate more on the patients feeling" after all it is the "ordinary everyday things in buildings that make an impact on people's lives" (Lawson, 2001).

Children's needs during their hospitalization process that are related to architectural design can be divided in two major categories. The first category is about their psychological balance which is related to the need of family's support, social interaction, security, territorial privacy, comfort, independence, access to outdoors, control etc. The second category is about the activities they can perform at a healthcare facility such as play (which is very important because according to Piaget children learn the world through this process), rest/sleep, eat, study, surgery/treatment, examination, visiting (Malkin, 1992). "...Architecture for children should adapt to the child's needs and scale, it is all about the child's perception of space and the way it can experience it. Spaces made for children should exceed the tiresome experiences of everyday life and consist of motivating factors that can awake the child's imagination and creativity" (Paraskeva, 2009). On the other hand when it comes to play and children it is not such a simple design task. "The institutionalization of play into formal times and places is thus an imposition on the child by adults, and is mainly done for the convenience of those adults. It is much easier to design environments that work well for children" (Lawson, 2001).

According to environmental psychology and a study that was preformed by Evans and McCoy in 1998 identified and discussed four aspects of the build environment (at the scale of a building) that can potentially influence human health and alter stress levels in individuals. Those aspects are stimulation, coherence, affordances and control (Evans, McCoy, 1998). This classification though, does not include the conditions that are fundamentally potential sources of physical stress. Among the conditions and factors that are potential causes of physical stress two basic categories have been identified: first the microclimatic conditions of a building and second the ergonomic conditions of use of the space (Del Nord, 2006). According to Piaget children need balance during all their developmental stages, so in the case of healthcare facilities' design for children the final result should be a space that offers the right amount of stimulation.

2.2 Current issues that affect hospital architecture

The field of architecture for health is at a turning point. Hospital architecture was immediately affected by global economy recession that begun in 2008. Hospitals nowadays are among the institutions suffering from the global economic malaise and many are reacting by laying off staff, limiting certain services and delaying or even cancelling construction projects (Verderber, 2010). The high costs of construction or renovation prevent private investors from financing.

Moreover, the issue of global climate change has a direct effect on hospital building architecture. The goal is carbon neutrality and regenerative strategies on new constructions and as much as adaptive reuse, preservation and increasing social capital through thoughtful environmental design (Verderber, 2010). The changing priorities of our current situation narrate some underlying trends for the future. The hospital needs reconsideration both as an idea and as a building type (Verderber, 2010). Last but not least, the issue of technology is rather crucial at this point. Nanotechnology, robotics, telemedicine, virtual reality and the rising role of nature in healing are matters that will influence the space and its layout.

2.3 Current trends of architecture in children's hospitals

The architectural work of hospitals, clinics, day hospitals or any other type of healthcare facility has always been a complicated task; even more nowadays that functionality is required to adapt to the new computer systems and sophisticated medical equipment must be combined with an aesthetics that is pleasant for patients, visitors and medical staff. All of the above has to be united in a modern, pleasant and welcome building that will offer safe and secure medical care. Sustainable design, eco-friendly materials and the emerging role of the landscape in the designing process are some of the most important considerations in healthcare facility design.

In architectural terms, modern healthcare facilities for children use the basic elements of design (such as exterior design, façades, natural environment, scale, colors, materials, light and shadows, movement and orientation) in an innovative and creative way along with the newest technological findings. "...To design a children's hospital is to listen to the many voices of children, ..., to their parents, to the staff, as well as to hear the advice of the environmental psychologist, the color specialist, the play and the music therapist, the pediatrician, the nurse, the intensivist, the social worker..." (Cleper-Borkovi, 2009).

Children are highly depended upon adults for comfort, guidance, and support, therefore during their hospitalization they need "...the presence of caring and attentive caregivers as well as their own parents" (Newman, 2009). In that case, some of the great challenges that rise are in relevance with the space in the patient's private room, or ways to create positive distractions that can take the mind (both the child's and the parent's) away from the pain or the stress. "The hospitals ward needs to reassure the patient who may be a little anxious, in some pain and later on a little bored too. Above all, hospital spaces need to counteract the loss of independence and identity the patient feels" (Lawson, 2001). The private room should be a pain-free space that the child and its family members can feel comfortable and support the child's emotional needs. Another very important factor that has an impact on the children's healthcare design is the special needs, habits and preferences that each one of the seven age groups during childhood and adolescence has. Starting from foetus, neonate, infant, preschool, secondary school, transition to adulthood, each stage has each own special characteristics, and each own scale and ergonometric features; therefore the healthcare environment has to adapt in order to provide safety and security to everyone. Furthermore, children during their hospitalization process have certain rights since the UN convention (1989) on the Rights of The Child and the foundation of the EACH (European Association for Children in Hospitals) Charter. Some goals of the EACH charter which are related to the layout and the design of the space are: the right of children to have their parents with them in the hospital, the possibility for children to have their own say in the care plan, opportunities for play and education in the ward, have contact with peers and a healing environment.

In order to clarify the different characteristics of design elements in children's healthcare facilities, it is preferred to separate them in categories; and further on study them through specific examples of contemporary architecture from around the world. Some of the most important elements of design are:

1. the exterior: The outcomes of contemporary architectural trends in hospital buildings aren't the massive buildings of the past that invaded their surroundings. The exterior of a children's healthcare facility is of great importance. "What children experience at the front door of a hospital, will color the impression of their entire stay". A children's hospital should be "simple and easily understood from the exterior" (Mead, 2005).

- 2. the interior (e.g. lobby reception and patient rooms): The smooth transition from an exterior space to an interior and the opposite is crucial in the design of healthcare facilities because it can affect both patients and visitors. Spatial organization of healthcare facilities relies on the continuity of spaces. The main entrance of a hospital building is always a prominent space, "...lobbies are the heart of a hospital welcoming all who came there and making an indelible and uplifting first impression" (Plappert, Gabel, Clements, 2005) it is a fact that the first impression of a lobby is decisive for the patient's feeling of well-being.
- 3. the relationship of the building with the natural environment: The bond between nature and well-being is not something new. It is known since the primitives societies existed. Local climate and building orientation specify the main guidelines of the design. The right position to the sunny side and to the main direction of the wind is of great importance for the patient's healing process. One of the basic criteria of an effective healing environment is also direct contact with nature. "...The hospital park, considered as a relaxing place for rest and entertainment for patients as well as an open space to public can play an essential role. The intention is that pain and isolation feelings are not the dominant ones" (Vavili, Stravela, 2009). Well designed exterior spaces like small parks, healing gardens, atriums and other open air spaces are likely to encourage social interactions.
- 4. the lighting (both natural and artificial): Lighting is an extremely powerful element in design. It is also an effective parameter for the patient as a powerful regulator of the body and its daily functions. The quality of lighting in hospitals can positively affect the emotional and medical state of a patient; it can also support well being and stimulate recovery (Antonakaki, 2009)
- 5. the colors and the materials used: The selection of materials and colors of the interior is always related to the use of the space and the people who will be using it. Architects who design and plan healthcare facilities know that "healing qualities depend on the choice of materials and colours" (Eggen, 2009). The proper choice can create an effective impact on the patient or the visitor. "Colors should offer warmth and variety without shouting or offending. Textures should be varied, inviting children to touch and learn about materials" (Mead, 2005). But colors and materials have to reinforce the functionality of the space and always support the patients.
- 6. the scale: the design must respond to all the child's needs and they should be adaptable for all the childhood's and adolescence's variations. For example low windows that can offer inviting views or trees without climbing on a chair or being picked up by an adult prevent children from feeling "imprisoned" in their hospital room (Filippazzi, 2009) The right scale can offer to offer to hospitalized children the sense of control of an unknown and strange environment to them. Also, "It is very important to introduce elements scaled to children of different ages and cultures, offering positive distractions" (Plappert, Gabel, Clements, 2005).

It is impossible to know the needs of the future hospital but what the real challenge right now for architects " is to build hospitals that reflect how hospitals actually work, not how we think they work and in ways that can be changed easily when needed" (McKee, 2010).

3 Proposed methodology

This dissertation is about the design of children's healthcare facilities and the way that architectural elements can create a more secure and less traumatic experience to hospitalized children; The importance of the environment during the hospitalization procedure and the resulting quality of the space.

The purpose of this research is to identify the current trends of architecture in this field, estimate the quality of children's hospitals and children's wards in the public general hospitals of Greece and propose interventions or redesign of the existing facilities that focus on a health promoting environment. Finally, there will be at least one design proposal of a case study of an adolescent's psychiatry unit in Greece.

The proposed methodology of the dissertation apart from the literature review through which the basic background information will be retrieved will consist of:

- Visual observation at the major children's hospitals and children's wards of Greece in order to collect data (about preferences of the local population regarding the use of space, the local education and culture about hospitals) that will be used to support a design proposal of the case study of an adolescent's psychiatry unit.
- 2. Questionnaire survey to staff, hospitalized children and parents at the major children's hospitals and children's wards of Greece that will give through statistical analysis the estimations about the quality of the local children's healthcare facilities. The results will be used in order to review and criticize the current Greek hospital building specifications and regulations.
- 3. Diaries, focus groups and in-depth interviews with hospitalized children for data collection which will give estimations about the children's needs during the hospitalization process and evaluation of the local healthcare facilities.
- 4. Case studies of up to date children's hospitals that will help in specifying the current trends of architecture in this area which will first of all support the design proposal of the case study of an adolescent's psychiatry unit.

5. Conclusions

The conclusions that can be predicted in advance are mainly about the proposed methodology. The evaluation of the most advanced case studies of children's healthcare facilities, will hopefully lead to notable proposals about the upgrading and improvement of the existing children's healthcare facilities of Greece. The data that will be gathered and statistically analyzed will lead to the evaluation of the current Greek hospital building specifications and regulations. The key aim of this research is to identify the design parameters that can contribute to the creation of a pleasant atmosphere, a healing environment and thus a better quality in children's healthcare

facilities; it also aims on the development and the upgrading of the quality of children's healthcare facilities in Greece.

Architectural design from that point of view is a powerful tool but can it really make the feelings of pain, fear and isolation during hospitalization process not the dominant ones?

References

Andrade, C.A., Lima, M.L., Fornara, F.F., Bonaiuto, M.B. (2012) User's views of hospital environmental quality: Validation of the Perceived Hospital Environment Quality Indicators, *Journal of Environmental Psychology*, 32, 97-111

Antonakaki, T.A. (2009) Aspects of Healing Environments, ZITI Publications, Thessaloniki

Aujoulat, I.A., Simonelli, F.S., Deccache, A.D. (2006) Health promotion needs of children and adolescents in hospitals: A review, *Patient Education and Counseling*, 61, 23-32

Bromley, E.B. (2012) Building patient-centeredness: Hospital design as an interpretive act, *Social Science & Medicine*, 75, 1057-1066

Cama, R.C. (2009) Evidence-Based Healthcare Design, Wiley, USA

Carayon, P.C (2010) Human factors in patient safety as an innovation, *Applied Ergonomics*, 41, 657-665

Cleper-Borkovi, F.C. (2009) Children's Hospital's Explorers (CHEX) 10 Issues and Innovations for Children's Hospital Planning and Design – The Language of Flowers – Thoughts on Designing a Children's Hospital, Proceedings of 28th International Public Health Seminar, Alinea editrice, Firenze

Coull, A.C. (2009) Children's Hospital's Explorers (CHEX) 10 Issues and Innovations for Children's Hospital Planning and Design – Thoughts on Designing a Children's Hospital, Proceedings of 28th International Public Health Seminar, Alinea editrice, Firenze.

Del Nord, R.D. (2006) Environmental Stress Prevention In Children's Hospital Design, Motta Architettura, Milano

Eggen, H.E. (2009) Aspects of Healing Environments, ZITI Publications, Thessaloni-ki

Evans, G.E., McCoy, J.M. (1998), WHEN BUILDINGS DON'T WORK: THE ROLE OF ARCHITECTURE IN HUMAN HEALTH. *Journal of Environmental Psychology*, Academic Press., 18, ps980089, 85-94

Filippazzi, G.F. 2009 *Also Walls Speak*..., The Culture For The Future Of Healthcare Architecture, Proceedings of 28th International Public Health Seminar, Alinea editrice, Firenze

Ghazali, R.G., Abbas, M.A. (2012) Newly Built Public Paediatric Wards Increase Length of Stay (LOS)?, *Procedia – Social and Behavioral Sciences*, 50, 623-632

Haron, S.H., Hamid, Y.H., Talib, A.T. (2012) Towards Healthcare Service Quality: An Understanding of the Usability Concept in Healthcare Design, *Procedia - Social and Behavioral Sciences*, 42, 63 – 73

Lawson, B.L. (1980) How Designers think the design process demystified, Architectural Press, London

Lawson, B.L. (2001) The Language of Space, Architectural Press, Great Britain

Malkin, J.M. (1992) Hospital Interior Architecture, creating healing environments for special patient populations, Van Nostrand Reinhold, New York

Mead, W.M. (2005) *Designing the World's Best Hospitals*, images publishing, Australia

McKee, M.M. (2010) *Translating service needs into capital solutions*, Proceedings of European Congress Healthcare Planning and Design (DuCHA), TNO, 6-9 June 2010, Rotterdam

Newman, C.N. (2009) Children's Hospital's Explorers (CHEX) 10 Issues and Innovations for Children's Hospital Planning and Design - Designing a Hospital for Children, Proceedings of 28th International Public Health Seminar, Alinea editrice, Firenze

Paraskeva, P.P. (2009) *Creative Spaces for Children*, Proceedings of 1st Pan-Hellenic Conference of Art and Environmental Art Education, Foundation of Secondary Degree for Education and Evgenidio Institution, Athens

Plappert, J.P., Gabel, L.G., Clements, D.C. (2005) *Designing the World's Best Hospitals*, images publishing, Australia

Samah, S.Z., Othman, N.O., Wahab, W.M. (2012) Assessing quality design of interiors: a case study of a hospital outpatient unit in Malaysia, *Procedia – Social and Behavioral Sciences*, 35, 245-252

Thakur, R.S., Hsu, SH, Fontenot, G.F. (2012) Innovations in healthcare: Issues and future trends, *Journal of Business Research*, 65, 562-569

Utkan, M.U. (2012) Children hospital design in children picture, *Procedia – Social and Behavioral Sciences*, 51, 110-114

Vavili, F.V. (2000) Children in Hospital: A Design Question, *World Hospital and Health Services*, 36, 2, 31-39

Vavili, F. V. (2009) Arts and Health Facilities, ZITI Publications, Thessaloniki

Vavili, FV, Stravela, S.S. (2009) Aspects of healing Environments, ZITI Publications, Thessaloniki

Verderber, S. V. (2010) Innovations in Hospital Architecture, Routledge, New York

A systematic literature review on the correlations of cognition and everyday functionality assessment tools for healthy older adults.

George Pavlidis1 and Ana Vivas2

¹ South East European Research Centre, Sociological Studies

gipavlidis@seerc.org

² International Faculty of the University of Sheffield, Psychology Department, City College

vivas@city.academic.gr

Abstract. The demographic increase of elders' representation in the general population poses certain social challenges and the need towards a healthy and active ageing process. An intact cognitive state is indispensable for the proper conduct of everyday activities in older age. The available research on the cognitive correlates of everyday functionality in older age however, has focused mainly on pathological aging, living the healthy spectrum of cognition largely unexplored. Most assessment tools of cognition and everyday functionality were primarily developed to examine clinical conditions, and therefore they may be less fit to examine the non-clinical population. This systematic literature review aimed to examine the cognitive domains, as well as the assessment tools for cognition and everyday functionality, that are most suitable for examining healthy seniors. The results indicate that general cognition, memory, and executive function seem to correlate frequently with everyday functionality in healthy seniors. The MMSE, TICS, and HVLT, seem to be the neuropsychological tests that correlate most frequently with everyday functionality in healthy seniors. In addition, the EPT and OTDL seem to be the assessment tools of everyday functionality that correlate most frequently with cognitive performance in healthy older adults. Some recommendations for further examination on the topic are discussed.

Keywords: cognition, everyday functionality, elders, healthy ageing, active ageing

1 Introduction

Advances in medicine, hygiene, and nutrition have resulted to a decrease in life threatening diseases and consequently to an increase of average life expectancy in western societies (Hamilton, 2011; Hertzog, Kramer, Wilson, & Lindenberger, 2009). This is a successful outcome: in conjunction however with the contemporary decline in birth rates, it fashions the ever growing of seniors' representation in the general

population (Atchey & Barush, 2004; Hamilton, 2012). Those older than 60 years are estimated to account for 20 % of the population in 2025 and at least 30% in 2060 (FutureAge, 2011; U.S Bureau of the Census, 1996 as cited in Atchley & Barush, 2004).

This demographic change into an ageing population poses certain challenges. Cognitive decline and increased rates of disability or frailty commonly accompany old age. Ageing in such terms is considered of poor quality with great social consequences, since it leads to dependency from a caregiver, increased costs for health services and a worrisome working capital/retirement pension ratio (Goodman as cited in Hamilton, 2012; Hertzog et al., 2009). These social challenges have stimulated the European Commission to form urgent policy objectives to: i) increase independent functionality in old age, ii) protect older workers participation in paid or unpaid work, and iii) ensure the sustainable integration and inclusion of the elderly in society. These are often referred to as Healthy Ageing or Active Ageing acts, depending on their focus of reference (European Centre for the Development of Vocational Training [Cedefop], 2011; FutureAge, 2011).

There is a large battery of research on the key components for a healthy and active life in the "silver years". More precise, seniors' cognitive and physical health is repetitively coined as central factors for the proper conduct of everyday activities. Such activities are for instance managing finances, commuting, socializing and operating devices (for reviews see Drag & Bieliauskas, 2009; Hertzog et al., 2009). Yet, cognitive decline with ageing is a natural and an inescapable process, as supported by neuropsychological (Farias et al., 2009) and neuroanatomical studies (for a review see Kramer, Bherer, Colcombe, Dong, & Greenough, 2004). Consequently, some decrements in everyday functionality as a function of age are expected and considered as natural in older age (Hamilton, 2011).

Yet, cognitive decline is not a temporal event, neither it follows a predetermined route. In the contrary, there is a great variance and heterogeneity in the distribution of cognitive functionality among the older population. Similarly, seniors' capability to live independently or to work is not predetermined over time. Virtually, it is manifested in diverse trajectories among seniors. In this vein, research of the ageing process aims to shed light into the factors that modulate cognitive decline trajectories. (Ardila, 2007; Birren, & Shaie, 2006; Hertzog et al., 2009). To date, many contributing factors have been identified, explaining some but not all of the variance on cognitive and everyday functionality performance in seniors.

The objective of examining seniors' cognitive state meets an immediate challenge: cognition is not a living organ in one's body, rather a latent entity that is inferred and measured by its products. Furthermore, cognition is the sum of many cognitive processes, vital for perceiving the world, for thinking and acting upon external and internal stimulation. As such, cognition cannot be depicted nor examined directly in fill detail, but mainly can be examined by its' products (i.e., cognitive performance) in specific tasks (Kramer, Bherer, Colcombe, Dong, & Greenough, 2004).

Following this line of scientific inquest, many neuropsychological tests that assess cognitive performance in various domains have been developed. Many were for other purposes than to measure "typical ageing", as most were a collateral product of research efforts to the prevalence of neurodegenerative diseases, or to diagnose neuropsychological pathologies (Brayne and Calloway, 1988 as cited in Cullum et al., 2000). Despite that, the available literature on cognitive performance and ageing denotes several cognitive processes to be more susceptible to the detrimental effect of the ageing process. These are the speed of processing, inhibition, verbal fluency, coordination and working memory (Ciolcombe and Kramer, as cited in Kramer et al., 2004; Farias et al., 2009). All these processes (except speed of processing) have been grouped together under the more general term of *executive functions*, which refers to a set of processes necessary to achieve goal-directed behaviours, such as those entailed in seniors' everyday activity (Hertzog et al., 2009).

The examination of elders' everyday functionality is based in assessments tools for basic Activities of Daily Living (ADL) or Instrumental Activities of Daily Living (IADL) (Reppermund et al., 2011). ADL questionnaires typically include questions that query behaviours such as bathing, grooming, dressing, eating etc. The IADL questionnaires typically include questions that query more complex tasks of everyday life such as managing finance, medication handling and housekeeping. IADL behaviours are considered to require higher levels of cognitive ability and more cognitive resources than ADL behaviours (Bell-McGhinty, 2002; Farias et al., 2009). In the available literature, IADL's measures have been more susceptible in performance decrements due to cognitive decline and age (Park, 1997 as cited in Royall et al, 2007).

Some IADL assessment tools are utilized complementarily in the diagnosis of cognitive-related disorders (e.g. dementia), or in granting disability schemes (Royall et al., 2007). In practice, most tools that assess cognitive or everyday functionality decline were designed for, and tested on, those that are reaching or bridging the threshold of disability (Cullum et al., 2002). Therefore, they are adequate to distinguish between the fit and disabled, but to the largest extend they may not be sensitive enough to distinguish various levels of functionality or performance in the non-clinical spectrum of cognition and everyday functionality (Farias et al., 2011). In a comprehensive review, Moore, Palmer, Patterson and Jeste (2007) differentiated between specific everyday functionality tests to be used for clinical practise (or for research on clinical populations), and those to be used on healthy older adults. For the examination of the latter population, the authors pinpointed the Everyday Problem Test (EPT; Willis et al., 1993), the Independent Living Scale (ILS; Loeb, 1996) and the Observed Task of Daily Living (OTDL; Diehl et al., 1995), as the most suitable ones.

Some studies that focused on seniors' cognition suggest that an intact cognitive state, particularly in the domain of executive function, may be crucial for everyday functionally. For instance, in the Royal et al (2007) meta-analysis, general cognitive status, executive control and memory were the aspects that explained the most variance in everyday functionality in older adults. In this study specifically, the neuropsychological tests that explained the most variance in everyday functionality were general cognition screening tools, general cognition composites, the Harvard Verbal Learning Test (HVLT; memory), and the Executive Interview (executive function).

The examination and promotion of active and healthy ageing however cannot be blindly based in disorder-diagnosis assessment tools. If so, the examination of factors that prevent from, or buffer against, unfavourable cognitive decline trajectories in older age may be misleading. For example, being married or living with others has been linked strongly with favourable cognitive decline trajectories in longitudinal studies of men (Van Gelder, 2007). This finding has been interpreted as the protective effect of marriage against cognitive decline. However, Van Gelder employed the Mini Mental State Examination (MMSE) to measure cognitive performance, which is considered to be of poor sensitivity in the non-demented spectrum of cognition (Kramer et al., 2004). Hence, the results may reflect more the protective effect of marriage against cognitive impairment, rather than against cognitive decline in general.

In a similar vein, the unsuitability to assess accurately cognition and everyday functionality in their healthy spectrum with the available measurement tools, may explain the modest correlation of cognitive performance with everyday functionality status reported in the meta-analysis conducted by Royall et al. (2007). This modest correlation was argued to indicate that either cognitive performance is not strongly related to everyday functionality, or that many of the available neuropsychological and everyday functionality measurement tools fail to capture it. The authors however, did not report (or did not took into account) the number of clinical and non-clinical participants in the included studies. It is therefore unclear, if the modest correlation between everyday functionality and cognitive performance reported was a product of a biased sample towards neurological disorders, which may have weakened the relation between cognition and everyday functionality further.

The recent interest on healthy and active ageing mandates the comprehensive examination of healthy, community dwelling seniors and not solely those part of the clinical population. It is important to examine cognition and everyday functionality in old age using assessments tools known to be sensitive to the non-clinical spectrum of cognition and of everyday functionality. To date, there is limited research whether the existing everyday functionality measurement tools offer significant correlations with, get predicted of, or predict, cognitive performance in healthy older adults (Bell-MacGinty, 2002; Farias et al., 2009; Farias et al., 2011; Reppermund et al., 2011). Therefore, this study was set out to review the relevant literature on the cognitive correlates of everyday functionality in older age, in order to identify which cognitive domains relate to healthy elders' everyday functionality. Moreover, this review aims at identifying which assessments tool might be more suitable for examining healthy older individuals in cognition and everyday functionality.

2 Method

In order to meet the objectives of the study, a systematic review of the available literature was conducted. The review included only peer reviewed articles available in the University of Sheffield online databases (e.g., Scopus, EBSCO, Web of Knowledge, APA), as well as in Google Scholar. The database search included the time period from the Royal et al., 2007 meta-analysis's last date of inquiry (August of 2005) till the November of 2012. The keywords included in the database search were "community dwelling", "elderly", "cognitive test", and "instrumental activities of daily living". We used broad terms such as "community dwelling" as keywords to capture a wider scope of seniors' everyday living and to exclude studies that focus on immobility, impairment or frailty. Respectively, we used narrow terms such as "cognitive test" and "instrumental activities of daily living", as these terms appear repetitively in the literature to describe the assessment tools used to measure seniors 'cognitive performance and everyday functionality respectively.

The database search yielded 410 articles, which were further reviewed. As in the Royal et al., 2007 meta-analysis, articles that reported multivariate regressions and correlations were included. Contrarily, results with logistic regression analysis or analyses of variance (ANOVA) were excluded. The rationale behind this exclusion was that such study designs tend to dichotomize, group or restrict cognitive and everyday functionality measurement scores into two or three values. On the one hand, this approach provides more easily significant results of potentially small effects (Royal et al, 2007). On the other hand, this approach tends to examine more the reaching or bridging of certain disorders' threshold rather the fluctuation of cognition and everyday functionality in seniors' healthy spectrum on these dimensions. Further exclusion criteria were articles focused in a certain illnesses (psychological or pathological, e.g. heart failure, diabetes, MCI, dementia, tremor), because of the radical confounding effect that these conditions have in the relation between cognition and everyday functionality. Articles where only one variable was examined (i.e., only cognition or only everyday functionality) were also excluded, as the main focus of the study was to examine the relation of cognition and everyday functionality and not each separately. Non-experimental design studies, or case studies, were also excluded, as they did not offer any empirical quantitative data.

The overall cognitive health of the resulting sample was estimated based on the sum percentage of those reported to be cognitively impaired or unimpaired in the included studies. Only studies that their samples were of more than 50% cognitively intact were included in the analysis. To population-based surveys, where the percentage of those cognitively impaired was not reported, the know prevalence of dementia (i.e., 7%) was assigned (Prince, Bryce, Albanese, Wimo, et al. 2003). The resulting articles were then analyzed to detect the significant correlations between neuropsychological measurements of cognition and everyday functionality assessment tools, using a standard counting procedure.

Results

From the total, 12 articles satisfied the study's criteria and were further reviewed (Figure 1.). The total of participants in these reviewed studies were 11496 ($M_{age} = 75.82$), from which 95% where cognitively intact and 5% cognitively impaired. These studies were conducted in USA, France, Hong Kong, Brazil, and Australia. Overall, 16 different neuropsychological tests and 17 different everyday functionality assessment tools were used, often combined and calculated in composite scores. This diver-

sity in the reported results of this study map multiple combinations and correlation between cognitive and everyday functionality assessment tools. No such combination was repeated nor replicated in at least two studies.

Firstly, in order to identify which neuropsychological assessment tools had the most correlations with everyday functionality, all everyday functionality assessments tools were conceptualized to measure the same construct (i.e., everyday functionality; see Table 1). Subsequently, in order to identify which everyday functionality assessment tools had the most correlations with cognition, all neuropsychological assessment tools were conceptualized to measure the same latent variable (i.e., cognitive performance; see Table 2). Three different versions of the MMSE were indentified in the studies, and they were grouped together for the counting procedure. Similarly, 6 different versions of the IADL assessment tool were indentified in the studies, and they were grouped together for the counting procedure.

The neuropsychological measures that had the most frequently reported statistically significant correlations with everyday functionality were the MMSE, HVLT, and TICS (for the assessment tools abbreviations see Appendix). Moreover, the EPT, IADL and OTDL were the everyday functionality assessment tools that had most frequently, reported statistically significant correlations with cognitive performance.

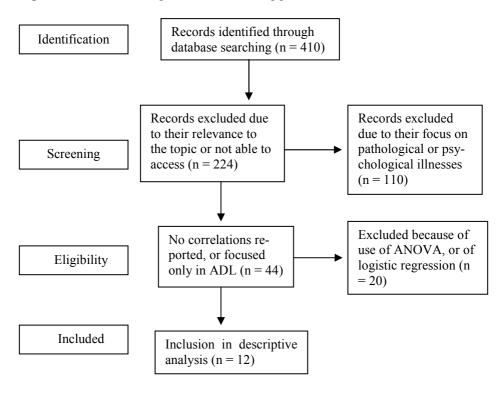


Figure 1. Prisma flow diagram on the screening process

Everyday functionality assessment tools	Neuropsychological assessment tools	Count
	MMSE - 3MS- cMMSE	5
	HVLT	4
	TICS	4
	TMT - A	3
	TMT - B	3
	AVLT	3
IADL (Lawton and Brody, Self or Informant report, COARS, Unclassified, Katz et al., Bayern), EPT, MWS,	Word Recall	3
	SDMT	3
	EXIT 25	2
SPPB, MDS-HC, HMS, IQCODE,	NART Language , Block design	2
DAD, OTDL, Direct Observation, DASF-R, KELS	Digit Span Back	1
	Letter Series	1
	VF	1
	Clox 1	1
	Clox 2	1
	Eckstorm's Vocabulary	1
	CRT	1
	Cognitive tests composite	1

Table 1. Frequency of reported correlations for everyday functionality with various neuropsychological tools in healthy older adults.

For the assessment tools abbreviations see Appendix

Table 2. Frequency of reported correlations cognition with various everyday functionality assessment tools, in healthy older participants.

	EPT	5
MMSE, 3MS, TICS, SDMT, EXIT 25, TMT- _{A.B} , Letter Series, Word Recall, Eckstorms Vocabulary, VF, Digit Span Back, HVLT, Clox 1.2, in the box, (NART)	IADL (all)	5
	OTDL	4
	KELS	3
	Direct Observation	3
	HMS	3
	MWS	2
	MDS-HC	2
Language, Block design	SPBB	1
	IQCODE	1
	DAD	1
	DASF - R	1

For the assessment tools abbreviations see the Appendix

3 Discussion

This study was set out to review the available literature on the cognitive correlates of everyday functionality in healthy older adults. Several articles were identified, where numerous correlations between cognition and everyday functionality were evident. These studies used several neuropsychological measures of cognition and assessment tools of everyday functionality. These findings replicate previous reviews (Drag & Bieliauskas, 2009; Hertzog et al., 2009; Royall et al., 2007) and highlight that cognition is related to everyday functionality, denoting however that this might be true specifically in the non-clinical proportion of the old as well. For the same part of the population, the findings indicate that general cognition, memory, and executive function correlate with everyday functionality. Consistent with the findings of Royall et al. (2007), some general cognition tests, the HVLT and some executive function tests (e.g., TMT A and B), have been most frequently reported to have statistically significant correlations with everyday functionality. Furthermore and consistent with the findings of Moore et al. (2007), the EPT and OTDL are the everyday functionality assessment tools that have been most frequently reported to have statistically significant correlations with cognition measurements in healthy older adults.

It seems the cognitive domains that relate to everyday functionality in healthy older adults are similar to those described previously in studies that did not focus specifically in the non-clinical proportion of the population (e.g., Royal et al, 2007). These cognitive domains are general cognition, memory and executive function. This was evident in neuropsychological measures of general cognition, using screening tools such as the MMSE and TICS, which although developed to detect the presence of dementia and MCI, offered repetitively significant correlations with everyday functionality on healthy adults. This is in odds with previous observations that claimed MMSE to be of poor sensitivity in the non-demented spectrum of cognition (Kramer et al., 2004).). An explanation might be the large mean age of the population included in this review ($M_{age} = 75.82$). The average life expectancy in the high-income countries is 80 years (World Health Organization, 2013). Taking into account the latter, the present analysis is based more to the older of old, an age range that in average has the lowest cognitive and everyday functionality scores. Therefore, MMSE and TICS may have been sensitive in the fluctuations of everyday functionality in this group, because they were used in individuals that in average are in greater risks for limitations and impairments in cognition and everyday functionality, a functionality range that MMSE and TICS were designed to capture. Thus, it is unclear whether MMSE and TICS correlate as greatly with everyday functionality in younger olds, where cognition and everyday functionality lie in higher levels of performance.

The HVLT, a broadly used neuropsychological test assesses verbal learning, delayed recognition, and memory, emerged as another frequent correlate of everyday functionality. This is consistent with the Royal et al (2007) findings, which placed HVLT to explain the most variance in everyday functionality in their reviewed studies. Similarly with the MMSE and TICS, HVLT seems to retain its relation to everyday functionality, even when putted in test specifically on healthy older individuals. The TMT- A and TMT – B tests that assess visual attention and task switching (the 484

latter is a component essential to executive function) also retain their relation to everyday functionality measures in healthy seniors. In extend, this may indicate them to be the appropriate neuropsychological tools to use in research that examines cognitive correlates of everyday functionality in healthy older adults.

In respect to everyday functionality assessment tools and their reference to the nonclinical population, the present findings indicate the EPT and OTDL are those that correlate most frequently with cognition in recent studies. Even that various IADL tests correlate evenly frequently with cognition, the variety of them (i.e., 7 distinct assessment tool of IADL) does not allow the distinction of one as more adequate for research purposes on healthy older individuals. The latter conclusion would be feasible only through a statistical meta-analysis of their effect sizes. The present study is restrictive to its descriptive nature and therefore, cannot inform of the overall strength between cognition and everyday functionality correlations, nor for the explained variance of the one to the other.

This is the most apparent limitation of the present study. The frequency with which several assessment tools of cognition and everyday functionality have been used in recent research and the frequency with which they offered significant correlations in the non-clinical senior population, cannot inform sufficiently for their adequacy in testing healthy older adults. It remains unclear why certain assessment tools were selected in each study, as the criteria for choosing them was vastly not reported (especially for cognition) in the reviewed studies. It is therefore unclear if those assessment tools of cognition and everyday functionality that have been reported more frequently to correlate with each other, are actually better tools in the examination of healthy seniors or if they just have been more popular for research purposes in these domains. In extend, this review cannot offer a conclusion for the comparative value among various assessment tools, rather it is mapping which instruments have been used in the past more frequently. In order to have more reliable indications of the assessment tools' sensitivity in the healthy spectrum of cognition and of everyday functionality, the relevant studies should be subjected to a statistical meta-analysis. Another important limitation of the present review is the participants' average age. More precise by, the participants' average age (76 years), compared to the overall average life expectancy in the world (70 years) and of the high-income countries (80; World Health Organization, 2013), indicates that the review refers more to the older of old. It is therefore unclear if this reviews conclusions can be replicated in studies that examine the correlates of cognition and everyday functionality in younger-old participants.

Future research should address these issues by conducting a thorough statistical meta-analysis on the cognitive correlates of everyday functionality in healthy, older adults. Moreover, cognition and everyday functionality, as well as the corresponding assessment tools, should be examined on young-old participants as well (e.g., between the ages 50 and 75).

Nonetheless, this study is a first step in identifying which cognitive domains correlate most frequently with everyday functionality specifically in healthy older adults. Furthermore, it provides some initial indications for the neuropsychological tests and everyday functionality assessment tools that retain their sensitivity in the healthy spectrum of cognition and of everyday functionality. The present findings indicate that tests assessing general cognition, verbal learning, delayed recognition, memory, visual attention and tasks switching, may relate to everyday functionality in healthy older adults. In this frame, the MMSE, TICS, and HVLT, are the neuropsychological tests that offered most frequently, significant correlations with everyday functionality of older adults in recent studies. Respectively, the EPT and OTDL are the everyday functionality assessment tools that correlated most frequently with cognition in recent studies of older adults. Consequently, and under our current knowledge, these assessment tools of cognition and everyday functionality may be the most adequate to research healthy older individuals, in the frame of advocating an active and healthy ageing process among the population.

References

Ardila, A. (2007). Normal aging increases cognitive heterogeneity: Analysis of dispersionin WAIS-III scores across age. *Archives of Clinical Neuropsychology*, 22, 1003-1011

Atchley, R.C., & Barush, A. (2003). Social forces and aging. An Introduction to social Gerontology (10^{th} edition), Wadsworth, Los Angeles

Bell-McGinty, S., Podell, K., Franzen, M., Baird, A., & Wlliams, M.J. (2002). Standard measures of executive function in predicting instrumental activities of daily living in older adults. *International Journal of Geriatric Psychiatry*, 17, 828-834

Birren, J.W., & Shaie, K.W. (2006). *Handbook of the psychology of ageing (6th edi-tion)*. Elsevier, Massachusetts

Cullum, S., Huppert, F.A., McGee, M., Dening, T., Ahmed, A., Paykel, E.S., et al. (2000). Decline across different domains of cognitive function in normal ageing; results of a longitudinal population-based study using camcog. *International Journal of geriatric Psychiatry*, 15, 853-862

Diehl M., Willis S.L., Schaie K.W (1995) Everyday problem solving in older adults: observational assessment and cognitive correlates. *Psychology and Aging*, 10, 478–91.

Drag, L.L., & Bieliauskas, L.A. (2009). Contemporary review 2009: Cognitive aging. *Journal of Geriatric Psychiatry and Neurology*, 23, 2, 75-93

European Centre for the Development of Vocational Training (2010). *Working and ageing: Emerging theories and empirical perspectives*. Publications Office of the European Union, Luxembourg

486

Farias, S.T., Cahn-Weiner, D.A., Harvey, D.J., Reed, B.R., Mungas, D., Kramer, J.H., et al., (2009). Longitudinal changes in memory and executive functioning are associated with longitudinal change in instrumental activities of daily living in older adults. *Clinical Neuropsychology*, 23, 3, 446-461

Farias, S.T., Mungas, D., Harvey, D., Simmons, A., Reed, B.R., DeCarli, C. (2011a). The measurement of everyday cognition: Development and validation of short form of the Everyday Cognition scales. *Alzheimer and Dementia*, *7*, 593-601

Farias, S.T., Mungas, D., Hinton, L., & Haan, M. (2011b). Demographic, neuropsychological and functional predictors of rate of longitudinal cognitive decline in Hispanic older adults. *American journal of Geriatric Psychiatry*, 19, 5, 440-450

FutureAge (2011). A road map for ageing research. Available at http://futurage.group.shef.ac.uk/road-map-flip-book.html

Hamilton, I.S (2012). *The psychology of ageing: An introduction*. Jessica Kinsley Publisher, London

Hamilton, I. S. (2011). Introduction. In Hamilton, I. S. (eds), *An introduction to gerontology*, Cambridge University Press, Edinburgh

Hertzog, C., Kramer, A.F., Wilson, R.S., &Linderberg, U. (2009). Enrichment effects on adult cognitive development: Can the functional capacity of older adults be preserved and enhanced? *Psychological science in the public interest*, 9, 1, 1-64

Kramer, A. F., Bhere, L., Colcombe, J. S., Dong, W., & Greenough, W.T (2004). Environmental influences on cognitive and brain plasticity during aging. *The journals of Gerontology*, 59A, 9, 940-957

Loeb P.A. (1996). Independent Living Scales (ILS). The Psychological Corporation, San Antonio

Moore D. J., Palmer, B. M., Patterson, T. L., & Jeste, D. V. (2007) A review of performance based measures of functional living skills. *Journal of Psychiatric Research*, 41, 97-118

Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., & Ferri, C. P. (2013). The global preference of dementia: a systematic review and metaanalysis. *Alzheimers Dementia*, 9, 1, 63-72

Reppermund, S., Sachdev, P.S., Crawford, J., Kochan, N.A., Slavin, M.J., Kang, K. et al. (2011). The relationship of neuropsychological function to instrumental activities of daily living in mild cognitive impairement. *Geriatric Psychiatry*, 26, 843-852

Royall, D. R., Lauterbach, E. C., Malloy, P., Coburn, K. L., Blak, K. J. (2007). The cognitive correlates of functional status: a review from the committee on research of the American neuropsychiatric association. *Journal of Neuropsychiatry and Clinical Neuroscience*, 19, 249-265

Van Gelder, B. M. (2007). Determinants of cognitive decline in older European men. Thesis. Wageningen University, Netherlands

Willis S.L, & Marsiske, M. (1993). Manual for the everyday problems test. University Park: Pennsylvania

World Health Organization (2013). Life expectancy. Available at http://www.who.int/gho/mortality_burden_disease/life_tables/situation_trends_text/en /index.html

Appendix

	Abbreviation
3MS	The modified MMSE
AVLT	Auditory Verbal Learning Test
MMSE	Mini Mental State Examination
CRT	Complex Reaction Time
CRT	Choice reaction time
DAD	Disability Assessment of Dementia Scale
DASF-R	Direct Assessment of Functional Status Scale
EPT	Everyday problem Test
HMS	Harvard Medical Schedule
HVLT	Hopkins Verbal Learning Test
IADL	Instrumental activities of daily lining
IQCODE	Informant Questionnaire of Cognitive Decline in the Elderly
KELS	Kohlman Evaluation of Living Skills
MDS-HC	IADL from Minimum-Dataset – Home Care
MWS	Maximum Walking Speed
NART	National Adult Reading Test
OTDL	Observed Tasks of Daily Living
SDMT	Symbol Digit Modalities Test
SPPB	Short Physical Performance Battery
TIADL	Timed IADL
TICS	The Telephone Interview for Cognitive Status
TMT	Trail Making Test
VF	Verbal Fluency

The development of executive functions in children

Fatbardha Qehaja-Osmani¹,

South East European Research Centre (SEERC), University of Sheffield ¹, fosmani@seerc.org

Abstract. Life success depends increasingly on the mastery of executive functions such as, working memory, shifting, and inhibiting. Executive functions, are cognitive processes that regulate behaviour, also influencing and predicting the development of other cognitive skills (Van der van et al, 2012). These functions are processes that regulate one's thoughts or behaviour. This paper will provide an overview of evidence regarding the development of executive functions in children, including key findings discussing their implications for children's learning, and indicating issues that merit further investigation.

Keywords: *Executive functions, cognitive flexibility, working memory, inhibition, children*

1 Introduction

Cognitive control mechanisms allow us to respond flexibly and efficiently to the changing demands of our environment, such as when a situation affords multiple possible responses (Snyder et al, 2008).

Executive function is a broad term that encompasses many higher order skills necessary for independent, goal-directed behaviour, including holding and manipulating information in working memory, switching between rules, inhibiting irrelevant information, planning, sequencing multistep tasks, and ascertaining the "big picture" from a complicated set of details (Denckla, 1989).

Executive functions play a crucial role in everyday cognitive tasks, such as estimating time and distance when going on holidays, or even when encountering work problems by planning solutions and implementing them.

The study of executive functions has been a popular field of study for several decades. Many studies (see Carreti, et al, 2004; Chyrsochoou & Bablekou, 2011) found relationships between executive functions and cognitive skills and processes important for educational achievement (e.g. comprehension skills and processes, visual processes, attention skills, etc.) which possibly results in bad or poor performance at school. Therefore, this paper has focused primarily on three EFs: working memory, cognitive flexibility (switching flexibly between tasks), and inhibition (deliberate overriding of dominant or prepotent responses) (see Miyake & Friedman, 2000).

2 Inhibition

Based on the theoretical model of Barkley (1997), inhibitory control is central to effective executive functioning in general. Inhibition is a mechanism that allows children to stop the activation of the current task, and switch to a new task (e.g. "OK enough with sorting cards based on their shape. Now we are going to match the cards based on the colour. So the red card goes with the red box and so on"). Inhibitory control develops significantly during the preschool.

In the *day-night task*, used to assess inhibition, children are presented with either a day picture of the sun or a night picture of the moon (and some stars) and are instructed to say 'night' to the day picture and 'day' to the night picture (see Gerstadt et al., 1994; Diamond et al, 2002). In this task the prepotent response of naming the stimulus must be suppressed in order to produce the 'conflicting' response required by the rule (Simpson & Riggs, 2005).

Similarly, Levin and colleagues (1991), proposed go/no go test in detecting the ability to inhibit responses. Levin and colleagues found that the greatest results in missed responses were found between the youngest (7-8) and the middle (9-10years) age groups with little or no further improvements in the oldest age group (13-15 years).

3 Working Memory

Working memory plays an important role in everyday cognitive task such as deciding which items to buy in a supermarket. Working Memory (WM) capacity reflects the amount of information that can be retained, while processing is carried out, for a short period of time, (Cowan et al, 2003). However the short term memory (STM) refers to the ability to temporarily maintain and store information.

Working Memory tasks, also called complex span tasks, are designed to evaluate the capacity of working memory by requiring both the maintenance of information for subsequent recall and some concurrent processing (e.g. reading sentences loud while trying to remember the last work of each sentence for a coming task)For example, in a counting span task (Case, Kurland, & Goldberg, 1982), children are asked to count dots on series of cards of ascending length and to recall the totals at the end of the series.

In an attempt to explore the effect of working memory and short term memory in cognitive flexibility in children, two groups of children are identified: children who could be perserverators or switchers. Perserverators are children who, despite been asked to move to a new task, they keep repeating the previous task which is not relevant anymore. Switchers are children who are able to switch to new tasks, as required, without repeating the previously relevant task. Thus, some researchers argue that switchers may have stronger working memory strength than perserverators, with stronger rule representation supporting both flexible switching and faster responses to switching (Blackwell et al, 2009). However, which executive functions are directly associated with failures in cognitive flexibility, still remains an important question for future investigations.

Diamond (1988) postulated that inhibition and working memory are inextricably linked, where if an individual is not able to maintain information over time and/or inhibit prepotent responses, he or she will continue to inflexibly choose the incorrect response. Furthermore, in recent studies when assessing cognitive flexibility in children, inhibitory control and working memory are tested jointly.

Similar to Barkley's theory Roberts and Pennington (1996) propose that normal development of executive functioning should be described in terms of the development of the two fundamental cognitive processes: inhibition and working memory. This was supported by the study conducted by Borella, Caretti & Pelgrina (2010), in which they posit that children with lack of inhibitory mechanisms have difficulty in controlling irrelevant information in the text to prevent overloading of memory capacity, which leads to damaging of the maintenance of relevant information and thus its integration as well as the comprehension and memorization of text.

4 Cognitive flexibility

People are often exposed to situations where they need to respond flexibly or even quickly e.g. after being used to cross the road in UK where people drive on the left side of the road, crossing the road in Greece where people drive on the right side of the road. When people are not able to flexibly switch to new demanding tasks they often perseverate or repeat older irrelevant behaviour or response. Children are particularly robust perseverators (Yerys & Munakata, 2006). In a DCCS (Dimension Change Card Sort) task pre-schoolers perseverate by continuing to sort pictures on cards according to their shape even when asked to switch and sort cards according to their colour (e.g. Zelazo, Frye, & Rapus, 1996). This perseveration occurs despite children being able to correctly verbally report what rule they should be sorting by.

In the course of their learning, children are often required to switch sources of information as task relevance changes. The study of task switching has become one of the major research tools in studying the dynamic and flexible control of task sets (Koch et al, 2010). In recent studies (Posner & Rothbart, 2007) findings show that switching is subserved by two distinct skills: selective attention and attention switching, sometimes called shifting. Attention switching enables a child to switch from one task to another as required by others, e.g. "let's move on to retelling the story", however in some cases children need to discover when a given rule has become invalid and switch to a new rule e.g. a child might be shown two items at a time and asked to indicate which one is correct.

Switching in preschool children is often assessed by the DCCS task. To date, the main finding is that although three year old children are able to follow instructions and sort without error on pre-switch trials, they have trouble switching sorting rules, but instead they tend to perseveratively sort by the pre-switch dimension (Hanania & Smith, 2009).

In this task, the child is not told which dimension is relevant but is asked to discover the relevant dimension through trial and error (Hannania & Smith, 2009). Recent studies claim that successful set shifting requires (a) determining which task goal is relevant and, if it is different from the previous one, (b) implementing a switch in task set (i.e., reorienting attention to the newly relevant information, selecting the correct response, and so on)(Craig, Chavalier, 2009).

An important current question is how these executive functions underpin cognitive flexibility, which is the ability to flexibly switch between routine behaviours or tasks (Munakata et al, 2012). Flexibility is often required when children are asked to treat a single object in two different ways (sorting objects by shape or sorting by color) or even when they themselves decide to change activities e.g. from hide-and-seek to playing cards. When they decide not to play the first game ("We're not going to play the hide-and-see game anymore") then the inhibitory mechanism needs to suppress the activation of this game in order for them to switch to a new game (Now we're going to play the card game"). During their card game they need to remember who was the last person who took the card, in this case they need to memorize and update previous information.

Executive functions have been considered to play a crucial role in bilinguals. Managing two languages at the same time requires a control mechanism that resolves the competition between two languages effectively in order to use one language; the target language (Poarch et al, 2012).

Modern research has supported the early findings (Costa et al, 1998) that when bilinguals use one language, either the target or non-target language, both languages are active and may compete for selection. Inhibitory control mechanism that enables the selection of languages, as proposed by Bialystok (2004), has larger impact on cognitive control in bilinguals over monolinguals.

5 Conclusion- Indications for future research

We cannot evaluate or investigate children's educational achievement, without relating it to their cognitive system and executive functions and the extent to which they affect children's lives. Several researchers have theorized that inhibition, as well as working memory and cognitive flexibility have a crucial role in children's educational performance.

Clearly more research is needed to better understand the nature of inhibitory mechanisms, working memory and short term memory as well as cognitive flexibility in young children. Furthermore, this literature review indicates that future research is needed to identify the impact that each function has in perseveration. Also concentration on the measurement of timing in WM processes indicates a need for future research (see Cowan et al, 2003). An important goal for future research will be to disentangle whether various forms of processing conflicts in task switching (i.e. at the stimulus level vs. at the response level, see Koch et al, 2010). Moreover, it would be interesting to further investigate the impact of bilingualism in all three executive functions.

The following questions are considered for future research: Does speaking a second language provide a benefit in children? Are bilingual children, compared to monolingual, faster switchers? Do monolinguals suffer less task interference than bilinguals, who are used to switching between two languages and inhibiting the one that is not currently being used in everyday life?

Answers to these and many other questions seem to constitute relevant research necessary and valuable.

6 References

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of e perimental Psychology*: Section A, 49(1), 5-28.

Borella, E., Carretti, B., & Pelegrina, S. (2010). The specific role of inhibition in reading comprehension in good and poor comprehenders. *Journal of Learning Disabilities*, 43(6), 541-552.

Carretti, B., Cornoldi, C., De Beni, R., & Palladino, P. (2004). What happens to information to be suppressed in working - memory tasks? Short and long term effects. *The Quarterly Journal of Experimental Psychology:* Section A, 57(6), 1059-1084.

Cragg, L., & Chevalier, N. (2012). The processes underlying flexibility in childhood. *The Quarterly Journal of Experimental Psychology*, 65(2), 209-232.

Cutting, L., Materek, A., Cole, A., Levine, T & Mahone, E. (2009). Effects on Fluency, oral language, and executive function on reading comprehension performance. *The international Dyslexia Association*. Sage Publication.

Diamond, A., Carlson, S. M., & Beck, D. M. (2005). Preschool children's performance in task switching on the dimensional change card sort task: Separating the dimensions aids the ability to switch. Developmental Neuropsychology, 28(2), 689-729.

Espy, K. A., & Bull, R. (2005). Inhibitory processes in young children and individual variation in short-term memory. Developmental Neuropsychology, 28(2), 669-688.

Hanania, R. (2010). Two types of perseveration in the Dimension Change Card Sort task. *Journal of experimental child psychology*, 107(3), 325-336.

Hanania, R., & Smith, L. B. (2009). Selective attention and attention switching: Toward a unified developmental approach. *Developmental Science*. pp 1–14

Koch, I., & Gade, M., & Schuch, S., & Philipp, A. (2010). The role of inhibition in task switching: A review. *Psychonomic Bulletin & Review*, 2010, 17 (1), 1-14

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive psychology*, 41(1), 49-100.

Miyake, A., & Shah, P. (Eds.). (1999). Models of working memory: Mechanisms of active maintenance and executive control. Cambridge University Press.

Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences*, 5(7), 309–315

Yerys, B. E., & Munakata, Y. (2006). When labels hurt but novelty helps: children's perseveration and flexibility in a card-sorting task. *Child Development*, 77(6), 1589–1607.

Zelazo, P. D., Craik, F. I., & Booth, L. (2004). Executive function across the life span. *Acta psychologica*, 115(2), 167-183.

Patterns of Stressors Among Albanian University Students: The Case of the European University of Tirana

Rudina Shkullaku1, Erika Melonashi1

¹ Department of Education, European University of Tirana, rudina.shkullaku@uet.edu.al, erika.melonashi@uet.edu.al

Abstract. In accordance with the recent reforms of the Albanian educational system, counseling services have started to operate in universities. These new developments produced the need to investigate stressor patters among Albanian students. The aim of the present study was to assess the patterns of stressors in one of the largest universities in Albania, the European University of Tirana. More specifically the study investigated the patterns of academic, intrapersonal, interpersonal, and environmental stressors in the context of a. gender differences and b. levels of study (first, second, or third year students). For this purpose 150 students completed the Student Stress Survey; moreover, an open ended section enquired about additional stressors. Results showed no gender differences in all four categories of stressors. As regards patterns of stressors by level of study, results showed a tendency for a decreasing number of interpersonal and intrapersonal stressors with increasing levels of study. Conversely, there was a tendency to report an increasing number of academic stressors with increasing levels of study. Findings suggested that counselors should be especially attentive to interpersonal and intrapersonal problems with first year students, helping them to face their inner developmental challenges and adapt to the social demands of a new environment. On the other hand, academic issues become especially important for final year students, since they have to face the challenge of the next step in their career (e.g., pursue a job or further qualification). Finally, responses to the open ended section suggested the need to consider cultural specific stressors in Albania, in future research.

Keywords: *academic, intrapersonal, interpersonal, environmental stressors*

1 Introduction

Counseling services represent an important aspect of student supporting schemes in European and American Colleges and Universities (Rott, 2006). Studies suggest that students find counseling services very helpful especially in coping with academic stress,

but not only (Durodoye, Harris & Bolden, 2000). Skowron, Wester and Azen (2004) found that interpersonal issues and other stressors outside the academic environment are also often brought into the counseling room, in search of expert help. In this context several efforts have been made to identify the types of stressors most often experienced by college and university students (Abouserie, 1994; Dill & Henley, 1998; Edwards, Hershberger, Russell, & Markert, 2001; Hudd, 2000; Pengilly & Dowd, 2000; Ross, Niebling, & Heckert, 1999). The identification of particular stressors as well as age or gender patterns of distribution have important practical purposes for the counseling psychologists in academic settings (e.g., Wenwen & Guangdong, 2010). This information provides a very useful preliminary theoretical framework for use in their everyday work with students (Foster, 2010). According to Feltham and Dryden, (2010) this aspect becomes particularly important in cases when counseling services have been just introduced and the structure lacks a specific overall frame. For instance, in Albania counseling services in universities have been adopted just five years ago and currently only five out of 35 universities have a functional system (Shkullaku, 2013). Moreover, counselors report that they lack information on the types or patterns of problems actually faced by the typical Albanian student, an aspect which makes the counseling process even more difficult (Shkullaku, 2013). Conversely, research from other countries on student stress, is quite extensive and might provide a useful framework to investigate this aspect in Albania.

2 Literature Review

University students are subject to several types of stressors, mainly related to the entrance into a new developmental stage, i.e., young adulthood (e.g., Pillow, Lovett, & Hill, 2008). According to Erikson (1968) at this stage, young adults are faced with issues such as intimate relationships, choice of a career, moving away from home, getting a first job etc. Therefore the transition from high school to college or university is often accompanied by efforts to cope with the high stress levels in adaptive or maladaptive ways (e.g., alcohol drugs) (Millstein & Halpern-Felsher, 2002;Webb, Bray, Getz, & Adams, 2002; Young, Morales, McCabe, Boyd, & D'Arcy, 2005).

However, studies have demonstrated that not only these major life events (e.g., beginning of studies at the university, starting a new job etc.) but also daily hassles (e.g., time pressure, arguments with colleagues etc.) are important stressors among university students (Goff, 2009; Kanner, Coyne, Schaefer, Lazarus, 1981; Le May, 2011; Ong & Cheong, 2009; Ross et al., 1999).

Nonetheless, recent research has shifted the investigative framework from the major/minor stressor distinction towards a classification of stressors based on their source (e.g., Ong & Cheong, 2009; Sharma & Kaur, 2011). This type of classification is in line with the investigative purposes of the studies, i.e., a classification which distinguishes

between academic and other types of stressors (e.g., intrapersonal, interpersonal, or environmental) has important conceptual utility. More specifically the term *Academic stressors*, refers to those stressors arising mainly from academic activities. This category includes major life events such as transferring schools, or graduation and daily hassles such as increased class workload, time management, conflicts with peers etc. Other types of stressors include *Interpersonal stressors*, which result from interactions with other people and include major life events such as starting a new relationship or breaking up and daily hassles such as working with new people, conflicts with parents etc. A third category of stressors includes *Intrapersonal stressors* resulting from internal sources; this category includes major life events such as death of a close person or injury and daily hassles such as change in eating or sleeping habits. Finally, *Environmental stressors* result from problems with the immediate environment, including major life events such as a change in the living environment or daily hassles such as problems with the car (Le May, 2011; Ross et al., 1999).

Although there is some general agreement about the types of stressors reported by university students, there are suggestions that there might be different patterns especially in terms of gender or level of study (Lindsey, Lyons, Hendricks, Mead, Butler, 2011; Matheny, Ashby, & Cupp, 2005; Misra, McKean, West, & Russo, 2000; Ukeh, Aloh, & Kwahar, 2011). Research shows that women are the most frequent users of these services as compared to men, and the kind of issues that they bring into counseling is also different; for instance women are more likely to ask for help on interpersonal issues as compared to men (Charbonneau, Mezulis, & Hyde, 2009;Didham & Csiernik, 2006; Robitschek & Hershberger, 2005). Additionally, specific patterns of help-seeking have been reported, depending on whether the student is a freshman or a senior; for instance there is a greater tendency of first year students not only to ask for help, but also to report more academic, intrapersonal, and interpersonal problems, probably due to difficulties in adjusting with the new environment (Barry, Hudley, Kelly, & Cho, 2009; Hudley, Moschetti, Gonzalez, Cho, Barry, & Kelly, 2009).

The purpose of the present study was to explore the patterns of stressors faced by Albanian students, in the context of two other variables: gender and year of study. In line with research from other countries it was expected that women would report a significantly greater number of interpersonal stressors as compared to men, but the two genders would not differ in terms of academic, intrapersonal, or environmental stressors. Secondly, it was expected that the number of academic, interpersonal, intrapersonal, and environmental stressors would decrease with the increasing level of study.

3 Data and Methodology

Participants and Procedure

The present study was conducted in one of the largest Albanian Universities, the European University of Tirana. Convenience sampling was used, i.e., questionnaires were distributed to students who were having classes on the specific days (and time) during which the researcher visited the settings. Data were collected during a time period of ten days. Participants were briefed about the purpose of the study and their right to withdraw at any time without penalty. They also signed an informed consent to take part in the study. The anonymous questionnaires were completed in class.

The study reached out to 170 students. The response rate was 88.2%, i.e., 11.7 % of the students did not accept to become part of the study. Therefore, the final sample for analysis consisted of 150 students, 81 men (54%) and 69 women (46%). The age of the participants ranged from 19 to 30 years with a mean age of 21 years (SD = 2.25). In terms of level of study, 55 participants were first level students, 46 were second level students and 49 were third level students.

Measures

The measure used in the study was the Student Stress Survey (Ross et al., 1999). This survey was constructed based on two pre-existing surveys, the Student Stress Scale (Insel & Roth, 1985) and the Taylor Manifest Anxiety Scale (Taylor, 1953). The measure was translated into Albanian and back translated into English by certified translators proficient in both languages. Furthermore, prior to the actual study, a pilot study (five participants) was conducted in order to make sure that the questionnaires were comprehensible to the participants. No problems were encountered during this procedure.

The questionnaire is a self-report including 40 items that assess academic, interpersonal, intrapersonal, and environmental stressors. The participant had to check (by putting X in the specific box) if any of the mentioned issues had troubled them in the past six months. Academic stressors were assessed with eight items focusing on school-related activities. This category included items such as: Increased class workload, Lower grade than anticipated, Anticipation of graduation, Serious argument with instructor, Transferred schools etc. Interpersonal stressors were assessed with six items evaluating interactions with other people and include. Items included: Change in social activities, Roommate conflict, Work with people you don't know, Fight with boyfriend/girlfriend etc. Both subscales showed acceptable internal consistency, Chronbach's $\alpha = .74$ and $\alpha = .72$, respectively.

Intrapersonal stressors were assessed with 16 items evaluating internal sources of stress. Items included: *Change in sleeping habits, Change in eating habits, New responsibilities, Change in use of alcohol or drugs, Outstanding personal achievement, Decline in personal health, Change in religious beliefs* etc. Environmental stressors were assessed with ten items evaluating problems in the environment (except for academic problems). This category included the following items: *Vacations/breaks, Computer*

problems, Messy living conditions, Car trouble etc. Both subscales showed acceptable internal consistency, $\alpha = .78$ and $\alpha = .75$ respectively.

Each category was scored in terms of the number of items that were checked. More specifically, for Academic stressors, the scores would range from 0 (none of the items checked) to 8(all of the items checked). Similarly for Interpersonal stressors scores would range from 0 to 6, for Intrapersonal stressors from 0 to 16 and for Environmental stressors, from 0 to 10.

Apart from the assessment of specific stressors the measure also included information on gender, age, and level of study of participants. Moreover, at the end of the questionnaire, an open ended question was added for the specific purpose of the study. Considering that the questionnaire had not been validated in Albania, participants were asked to list any other situations which might have caused stress during the last six months within or outside of their academic environment.

4 Empirical Analysis

Considering the categorical nature of the data, nonparametric tests were conducted. Therefore Chi^2 tests were performed to assess gender differences in academic, intrapersonal, interpersonal, and environmental stressors. Results showed that Chi^2 for gender differences in academic, intrapersonal, interpersonal, and environmental stressors were not significant. Therefore, male and female students did not differ in the number of stressors they reported in each of the four categories (see Table 1).

Variable	M (SD)		Chi^2	p-values	
	Men	Women			
	(N = 80)	(<i>N</i> =67)			
1.Academic	2.86 (1.00)	2.7(1.20)	13.37	.20	
stressors					
2.Interpersonal	3.46 (.78)	3.19 (1.06)	7.87	.64	
stressors					
3.Intrapersonal	3.92 (1.31)	3.85 (1.30)	10.72	.55	
stressors					
4.Environmental	3.42 (.84)	2.47(1.03)	7.81	.80	
stressors					

Table 1. Chi² Tests for Differences between Male and Female Students in Academic, Interpersonal, Intrapersonal and Environmental Stressors

Spearman correlations were performed to investigate the relationships between the four categories of stressors and levels of study. Results indicated significant relationships between level of study and academic stressors, $\rho = .43$, p < .05. Therefore, students at an

earlier stage of their studies tended to report less academic stressors than those at a later stage. However, there were negative relationships between level of study and interpersonal, intrapersonal stressors, $\rho = -.42$, $\rho = -.49$, p < .01. Thus students at an earlier stage of their studies tended to report more interpersonal and intrapersonal stressors than those at a later stage (e.g., third year students). Finally there were no significant relationships between environmental stressors and levels of study, $\rho = -.11$, p > .05 (see Table 2).

Variable	M (SD)	1	2	3	4
1. Academic stressors	2.85 (1.10)				
2. Interpersonal stressors	3.34 (.91)	.01			
3. Intrapersonal stressors	3.39 (1.83)	11	.34**		
4. Environmental stressors	3.44 (.93)	.02	.29*	.17*	
5. Level of Study	-	.43*	42**	.49**	11

 Table 2. Spearman Correlations Between Level of Study and Academic, Interpersonal, Intrapersonal and Environmental Stressors.

Levels of significance: p < .05, p < .01

As regards the open-ended section, only 70 (out of 150) students filled in their comments. The main concerns reported by these students involved the academic, interpersonal, and intrapersonal domains. Stressors reported in the academic domain included difficulties to understand the new academic requirements, perceived inability to adapt to the higher education system, worry over the significance of the university degree in the prospective job market etc. Stressors reported in the interpersonal and intrapersonal domains included domestic violence, parental problems with alcohol (interpersonal), unexpected pregnancy, inability to control depressive/anxious moods (intrapersonal) etc.

5 Conclusions

The aim of the present study was to determine the patterns of reported stressors among students at the European University of Tirana. Differences were expected in terms of gender; more specifically it was hypothesized that male and female students would differ in their reports of interpersonal stressors but not intrapersonal, academic, or environmental stressors. Results showed no gender differences in any of the stressor categories. The second hypothesis suggested that the number of academic, interpersonal, intrapersonal and environmental stressors would decrease with increasing level of study. Results partially supported the hypothesis as negative relationships were found for the interpersonal and intrapersonal domains, but positive relationships were found for academic stressors. Finally, no significant relationships were found between levels of study and environmental stressors.

The absence of any significant gender differences across all categories of stressors suggested that apparently both genders are troubled by a similar number of academic, intrapersonal, interpersonal, and environmental stressors. Nonetheless the present findings do not imply that men and women experience similar levels of stress. Therefore one of the issues that need to be addressed in further research has to do with the intensity of the stress experienced. For instance, although both men and women might consider 'fighting with boyfriend/girlfriend' as a stressor, they might differ in the importance they place on it (E.g., women might rate it as 'extremely stressful', while men as 'moderately stressful'). Even so, it should be mentioned that the overall number of stressors reported for each gender and each category was moderate to low, which in itself represents a quite positive finding; nonetheless care should be taken since there is also the possibility that the items of the questionnaire did not actually represent the 'actual stressors' of Albanian students. Responses on open ended items suggest the need for future modification of the questionnaire, in order to adapt it to the specific context of the research. Moreover, the inclusion of gender specific stressors as suggested in the open-ended questions (e.g., pregnancy or domestic violence for women) represents an issue that should be carefully considered in future research.

Findings on a positive relationship between academic stressors and levels of study suggested the need to direct counseling attention not only towards first year students dealing with adaptation issues but also older students who are about to graduate. Issues such as the graduation itself or what to do next (career counseling) might be of particular interest to these students; the specific Albanian context of an exponentially growing number of graduates from public and private universities on the one hand and the scarcity of jobs on the other, make this kind of issues even more pressing. Indeed, as suggested in the open ended section 'the perceived value of the degree' in the actual job market seems to be an important issue for graduating students.

On the other hand, a focus on intrapersonal and interpersonal stressors might be especially important for first year students. The management of interpersonal relationships as well as self-management seem to be pressing issues at these earlier stages of the academic life. These findings are in line with studies suggesting adaptation and maturation effects (Erikson, 1968; Pillow et al., 2008; Ross et al., 1999). From a practical perspective, the results suggest that apart from individual level counseling, specific programs (group activities) might be designed which allow students to improve their interpersonal skills. Intrapersonal problems also need particular attention in these early years; indeed as suggested through the open-ended responses, inability to control mood changes (depressive or anxious) seem to be particularly troubling for at least some of

these students. Future research might examine further this particular aspect, e.g., by assessing the symptoms of these specific disorders (anxiety and depression).

In conclusion it must be mentioned that the present study also has several limitations. For instance, the findings cannot be generalized to other Albanian university students, apart from the European University. Furthermore, it is clear especially from the answers at the open ended section that future use of the scale needs adaptation for the specific Albanian context. Moreover, in order to make more solid conclusions the measure of stressors needs to be accompanied with a stress-assessment questionnaire evaluating the relevance of each stressor. Despite the several limitations the findings of the study provide a first step towards the identification of stressor patterns among Albanian students. Future research might address these same issues, but with improved measures and a larger representative population of Albanian students.

6 References

Abouserie, R (1994) Sources and levels of stress in relation to locus of control and selfesteem in university students. *Educational Psychology*, 14, 3, 323-332.

Allen, AD Coyne, JC Schaefer, C Lazarus, RS (1981) Comparison of two modes of stress measurement: Daily hassles and uplifts versus major life events. *Journal of Behavioral Medicine*, 4, 1, 1-39.

Barry, LM Hudley, C Kelly, M Cho, S (2009) Differences in self-reported disclosure of college experiences by first-generation college student status. *Adolescence*, 44, 173, 55-68.

Charbonneau, AM Mezulis, AH Hyde, JS (2009) Stress and emotional reactivity as explanations for gender differences in adolescents' depressive symptoms. *Journal of Youth and Adolescence*, 38, 1050-1058.

Didham, S Csiernik, R (2006) Gender matters in the counseling setting: Gender sensitive treatment can enhance women's gains. *Addiction Professional*, 4, 2, 16-42.

Dill, PL Henley, TB (1998) Stressors of college: a comparison of traditional and nontraditional students. *The Journal of Psychology*, 132, 1, 25-32.

Durodoye, BA Harris, HL Bolden, VL (2000) Personal counseling as a function of the community college counseling experience. *Community College Journal of Research and Practice*, 24, 6, 455-468.

Edwards, KJ Hershberger, PJ Russell, RK Market, RJ (2001) Stress, negative social exchange, and health symptoms in university students. *Journal of American College Health*, 50, 75–79.

Erikson, EH (1968) Identity, youth, and crisis, W. W. Norton Company, New York.

Feltham, C Dryden, (2010) *Brief Counseling: A practical integrative approach*, Open University Press, New York.

Foster, HL (2010) A best kept secret: Single-subject research design in counseling outcome research and evaluation. *Professional School Counseling*, 6, 2, 146-154.

Goff, AM (2009) Stressors, academic performance, and learned resourcefulness in baccalaureate nursing students. *International Journal of Nursing Education*, 8, 1, 147-159.

HUCS (2002) Survey into medical, psychiatric and counseling provision in higher education. Retrieved from http://www.hucs.org

Hudd, S Dumlao, J Erdmann-Sager, D Murray, D Phan, E Soukas, N Yokozuka, N (2000) Stress at college: Effects on health habits, health status and self-esteem. *College Student Journal*, 34, 2, 217–227.

Hudley, C Moschetti, R Gonzalez, A Cho, S Barry, L & Kelly, M (2009) College freshmen's perceptions of their high school experiences. *Journal of Advanced Academics*, 20, 438-471.

Insel, PM Roth, WT (1985) Core concepts in health (4th Ed.), CA: Mayfield, Palo Alto.

Kanner, AD Coyne, JC Schaefer, C Lazarus, RS (1981) Comparison of two modes of stress measurement: Daily hassles and uplifts versus major life events. *Journal of Behavioral Medicine*, 4, 1, 1-39.

Lago, C (2002) The university and the wider community. In Stanley, N Manthorpe, J. (Eds.) *Students' mental health needs: Problems and responses*, Jessica Kingsley Publishers, London.

Le May, M (2011) Stress trends at Georgetown University. An analysis of ACA-NCHA. Unpublished thesis: Georgetown University.

Lindsey, R Lyons, R Hendricks, D Mead, A Butler, LK (2011) Sources of stress: Gender and classification for African American college students. *College Student Journal*, 45, 4, 110-120.

Ong, B Cheong, KC (2009) Sources of stress among college students: The case of a credit transfer program. *College Student Journal*, 43, 4, 250-262.

Matheny, KB Ashby, JS Cup, P (2005) Gender differences in stress, coping, and illness among college students. *Journal of Individual Psychology*, 61, 4, 365-375.

Millstein, SG Halpern-Felsher, BL (2002) Judgments about risk and perceived invulnerability in adolescents and young adults. *Journal of Research on Adolescence*, 12, 399-422.

Misra, R McKean, M West, S Russo, T (2000) Academic stress of college students: Comparison of student and faculty perceptions. *College Student Journal*, 34, 236–245.

Mori, SC (2000) Addressing the mental health concerns of international students. *Journal of Counseling and Development*, 78, 137–144.

Pengilly, JW Thomas, ED (2000) Hardiness and social support as moderators of stress. *Journal of Clinical Psychology*, 56, 6, 813-820.

Pillow, HB Lovett, SB Hill, V (2008) Children's, adolescents', and adults' reference to goals to explain interpersonal actions. *Infant and Child Development*, 17, 310-317.

Ross, SE Niebling, BC & Heckert, TM (1999) Sources of stress among college students. *College Student Journal*, 33, 2, 312-317.

Rott, G (2006) The globalization in the higher education: Challenges and opportunities for practicing psychological counseling. *Review of Counseling and Guidance*, 56, 1, 97–106.

Robitschek, C Hershberger, RA (2005) Predicting expectations about counseling: Psychological factors and gender implications. *Journal of Counseling and Development*, 83, 4, 450-460.

Sharma, N Kaur, A (2011) Factors associated with stress among nursing students. *Nursing and Midwifery Research Journal*, 7, 1, 244-248.

Shkullaku. R (2013) *Psychological counseling is a new innovation and practice in higher education institutions in Albania.* Paper to be presented at the International Conference on Integration and Innovation in Research and Education, Prague, 2013.

Skowron, EA Wester, S Azen, R (2004) Differentiation of self, mediates college stress and adjustment. *Journal of Counseling and Development*, 82, 69-78.

Taylor, JA (1953) A personality scale of manifest anxiety. *Journal of Abnormal and Social Psychology*, 48, 285-290.

Ukeh, MI Aloh, PK Kwahar, N (2011) Stress and gender in relation to self-esteem of university business students. *Gender and Behavior*, 1, 210-215.

Webb, J Bray, J Getz, J Adams, G (2002) Gender, perceived parental monitoring, and behavioural adjustment: Influences on adolescent alcohol use. *American Journal of Orthopsychiatry*, 72, 392-400.

Wenwen, S Guangdong, W (2010) Research on mental health of college students (from 2007 to 2009) and analysis. Art and Literature for the Masses, 8, 157-189.

Young, AM Morales, M McCabe, SE Boyd, CJ D'Arcy, H (2005) Drinking like a guy: Frequent binge drinking among undergraduate women. *Substance Use and Misuse*, 40, 241-267.

The importance of children's participation in decisionmaking about themselves and the barriers, which might encounter.

Theodosios N. Tasios¹ and Efrosini Kalyva¹

¹South-East European Research Centre (SEERC), Greece

t.n.tasios@gmail.com

Abstract. All children, including those with special educational needs, as citizens of a country and as individuals have their own beliefs and the right to participate in decision-making about themselves, it has been established by the United Nations Convention on the Rights of the Child (UNCRC). The purpose of this study is not only to explain the meaning of the term "participation", many experts have examined the term "participation", such as Hart (1992), Shier (2001) ("pathways of communication"), Treseder (1997), but also to highlight the particular benefits of this process, along with the barriers which these children encounter when they try to exercise this right (Cavet & Sloper, 2009).

Keywords: children, decision-making, participation

1 Introduction

This paper outlines the importance of children's participation in decision-making about themselves, along with models of participation, which have been proposed in the past and the barriers which children might encounter. The paper suggests that while many authors proposed quite enough studies about models of participation and the benefits for children of that process, still now the meaning of the term "participation" is controversial and the adults set enough barriers to the participation of children.

2 Children's rights and regulations

Children as citizens of a country and as individuals have their own beliefs and the right to participate in decision-making about themselves. Adults have the tendency to

not take into consideration children's opinion, because they work on the assumption that younger children do not know what is right or wrong. However, according to the United Nations Convention of the Rights of the Child (UNCRC) in 1989 and more specifically in article 12, children have the right to participate in decisions which influence themselves. Wherever literature mentions children's rights, it refers to those human rights that specifically deal with the protection of children. The United Nations Convention on the Rights of the Child (UNCRC) is a comprehensive, internationally binding agreement on the rights of children, adopted by the United Nations (UN) General Assembly in 1989. It incorporates children's: civil and political rights (like their treatment under the law, part I); social, economic and cultural rights (like an adequate standard of living, part II) and protection rights (from abuse and exploitation, part III). Part I, Article 12, of UNCRC defined the children's rights to participation which are "a) State Parties shall assure to the child who is capable of forming his or her own views the right to express those views freely in all matters affecting the child, the views of the child being given due weight in accordance with the age and maturity of the child' and b) that the child shall in particular be provided the opportunity to be heard in any judicial and administrative proceedings affecting the child, either directly, or through a representative or appropriate body, in a manner consistent with the procedural rules of rational law". In Greece that convention was adopted in 1992 (Convention on the Rights of the Child, (FEK, N 2101 192/2-12-1992). It should be noted here that before the article 12 of the UNCRC, it was existed the "consumer movement" in the 1970 and in recent years that named "user involvement", which had purpose to promote the power along with the influencing of these children and make more services available for these people. Although, the previous movement influenced only UK Government's Modernising Agenda (Sinclair, 2004). By definition, all children have these rights including those with special educational needs. In addition, article 12 does not require adults to agree with the child's opinion, but to take into account when they make a decision which affects the child (Kramer Olsen, Mermelstein, Balcells, & Liljenquist, 2012; Saaltink, Mackinnon, Owen, & Tardif-Williams, 2012).

Also the Children Act guidance and regulations, which influence children with disabilities specifies, that these children cannot be assumed as incapable to participate in decision-making and according to Mental Capacity Act 2005, all people over 16 years old have the capacity for decision-making, unless it is proved that they lack capacity. Capacity defined as decision-specific for a specific time (Franklin & Sloper, 2009). In addition, specialists, such as psychologists, special educators, should protect and promote the participation of these children (Franklin & Sloper, 2009).

3 Definitions and models of participation

For understanding the importance of the child's opinion it should be explained what participation in decision making is. First of all, there is no single definition of participation, in ICF (International Classification of Functioning, Disability and Health) the term participation defined as "involvement in a life-situation" (World Health Organization, WHO, p.7). Definitions of the UNCRC and of the ICF are very broad, so participation can refer to a range of activities, such as children taking part in adult-orienting facilities, participation in formal and informal programmes or activities, participation in decision-making and many other forms, which all have different goals (Kramer, Olsen, Mermelstein, Balcells, & Liljenquist, 2012; Moses, 2008).

Coster and Khetani (2008) also noted a gap on the definition of the term participation on ICF. More specific, they distinguish participation from activity, because participation as "life situation" is the engagement in life contexts and on the other hand activity is an execution of a specific task or action. Moreover, according to Boyden and Ennew (1997) participation can just mean "taking part" / "being present", or the transfer of power which has as a result the views of participants to have real influence on decisions. So according to Sinclair (2004) participation is a term, which lack precision and is used to describe many different activities, which taking place in different circumstances. Still now there is debate about the exact definition of the term participation and how to empower children, especially these with disabilities, in order to participate more in decision-making. Although the first overviews about participation and the improvement of that were produced by R. Hart based on a modified version of Arnstein ladder of participation (Hart, 1992; Horelli, 1998).

Hart (1992), who argued that "participation in society begins from the moment a child enters the world and discovers the extent to which she is able to influence events by cries or movements" (p.4), proposed a model of participation (ladder of participation). More specific, according to Hart (1992) participation is "the process of sharing decisions which affect one's life and the life of the community in which one lives. It is the means by which democracy is built and it is a standard against which democracies should be measured." (p.5). So, it is the democratic right of individuals to be involved in decision-making and to influence those decisions which affect their. Hart's ladder of participation has 8 levels, in hierarchical order, each of these reflect the degree of authority and power that is shared between adults and children (Day, 2008; McNeish, 1999).The first three levels named by Hart as "*manipulation*", "*decoration*" and "*tokenism*". These three are often presented as a form of participation, but in fact are only *decoration* without actual participation of children (Hart, 1992; Head, 2011). As the child goes up a level, there is an increase in his or her participation. On the top are children who are able to initiate their own projects.

More specific, the lowest level of Hart's ladder has the name *manipulation* and refers to those cases where adults having awareness, use children's voices and actions for their purposes but children do not understand their actions. According to Hart a good example of children's *manipulation* is when children used from adults for political purposes, carrying political placards (Hart, 1992; McNeish, 1999).

McNeish (1999) analyzed one more level in Hart's model with the name "*deception*" and according to McNeish it is more common. *Deception* is used according to Hart (1999) when adults with good intensions deny their own involvement in a task with children, in order to think the other people that the task was done only by children. These adults supposed that if they referred to their involvement that will have the result to reduce their effectiveness (McNeish, 1999).

The next level after *manipulation* is *decoration*, where adults do not pretend that the project was motivated by children, but they just use the children to support their task in a non direct way (Hart, 1992; McNeish, 1999). An example which was used by Hart (1999) is when a child wear a T-shirt related to something and might dance or sing with that at an event, but he/she has no notion about the cause or the relation of the T-shirt with the occasion (Hart, 1992; McNeish, 1999).

The most common form of non-participation of children in adult's tasks, according to Hart (1999) is that with the name of *tokenism*. It is used for occasions where adults give voice to children, but in fact children have little or no opinion neither have the chance to make their own choice. A relative example, according to McNeish (1999), is when children go to a conference of adults because they must go, but they know that no of their ideas will be taken into consideration of adults, although they serve a symbolic action, in order to be sure the adults that the views of children will be taken into account, without any in depth and meaningful effort to do that.

The next five levels are, according to Hart (1999) forms of genuine participation. The first level of actual participation after *tokenism* is *Assigned but informed*, which refers to situations where adults define and assign the roles which will be undertaken by children. So, children are informed and understand their role in the certain activity, but it still remains a lower form of participation, because adults set the agenda and the limits according to which children must behave with the appropriate way (McNeish, 1999).

Some of the previous projects, which mentioned above could be moved up on the participatory ladder if the consultation was meaningful. The fifth rung of Hart's ladder has the name *Consulted and informed* and it refers in these activities where children work as consultants for adults. More specific, the specific ladder it refers in these projects where adults designed it, but children first of all understand the process, are consulted and finally their views are taking into serious account (Hart, 1992; McNeish, 1999).

The next step after consultation is to share decisions between adults and children and that is the reason why Hart named the next level of his ladder *Adult Initiated*, *Shared Decisions with Children*. It is very often, as McNeish (1999) noted, adults even after consultation of children to take the final decisions alone and that is the difference at that level, that children are involved in decision-making. In other words children at that level are involved at all parts of an activity, from the design to the evaluation and the execution (Hart, 1992; McNeish, 1999).

The semi-final and rare level of Hart's model is the *Child-Initiated and Child-Directed* level, where as the name indicates it refers in complex projects, which designed and directed by children without adults involvement (McNeish, 1999).

The highest rung of Hart's model has the name *Child-Initiated, Shared decisions with Adults* and it characterizes situations where children design and manage a project, but decide to share their decisions with adults. Although, that rung is very rare, neither because adults do not trust children, nor that they have the desire to help children, but also because it is great challenge for adults to not take over them the whole project.

Hart's model has been criticised by many authors, because it is a hierarchical model. More specific, according to Hart's ladder of participation, children and young people must be at the highest rung of the ladder in all of their activities in order to participate actively. Although, there are some activities, such as these contexts of political violence, where children must participate less than adults, in order to be protected. So, it would be better if the level of participation determined by the specific context, social relationships, cultures and groups, because different levels may be more appropriate for different activities (Day, 2008; Hammel et al, 2008; Hinton, 2008; Sinclair, 2004).

For that reasons Treseder (1997), rearranged Hart's participation rungs in a non linear and hierarchical model, but into a circular model (Mannion, 2003; Hinton, 2008). According to Treseder none of the levels is assumed to be better or to work better than another level (Kirby & Bryson, 2002). In addition, he noted that different levels of participation may be appropriate for different groups of children and young people (Day, 2008; Sinclair, 2004).

One other model, "*pathways of communication*", which is based on Hart's, has been introduced by Shier (2001) and consists of 5 levels, where as children step up a level there is an increase in their participation. The five levels are *children are listened to*, *children are supported in expressing their views, children's views are taken into account, children are involved in decision making-processes and children share power and responsibility for decision-making*. These levels are similar to the levels of participation of Hart's model, but Shier's model is different, because each level consists of 3 sub-levels, *openings, opportunities and obligations* (Day, 2008; Gray 2002; Horwath et al., 2012). An *opening* occurs when an individual have the willingness to actively involve other users. *Opportunity* refers to the availability of the appropriate skills knowledge and resources in order to enable the user to work. Finally, *obligations* reflects the agreed policy, the extent of the involvement and the expectation of that participation (Day, 2008; Gray 2002). According to Shier it is important to gradually involve children in decision-making and to support the mechanisms that increase children's control over decisions (Horwath et al., 2012).

Finally, Sinclair (2004) proposed another model of participation with four levels, where the first is again similar with Hart's and the others are, *focus of decision*-

making (public or private and individual or collective user, because the mechanism for the involvement of the children is different), although, as demonstrated in some studies the participation of children into family context and into social or public domain are intricately linked (Moses, 2008), nature of the participation activity (one-off consultation event or on-going involvement, for example forums, involvement in youth councils or organization or in decision-making and which affect the term children, according to the UNCRC all those who are under 18 years old) and characteristics of the children and young people involved (social, cultural and mental health needs) (Day, 2008). Although, except these authors many others proposed almost similar models, such as Barker (1999) zones penetration, Lansdown (2001) and others.

4 Benefits of participation

The interest for the level of participation of children in decision making together with the factors which affect participation can be explained by the many positive results which participation might have. Many authors highlighted the positive outcomes, which might have the participation of children in decision-making (Coyne, & Gallagher, 2011). First of all it has been proved that children with typical development, who were engaged into meaningful participation, by making plans or decisions in their preschool age, later at the age of 19 they are less involved in drug use, teen pregnancy, delinquency and school failure, but in contrast they have higher earnings (Benard, 1993; Rosenberg, Jarus, Bart, & Ratzon, 2011).

The benefits of participation also noted by Day (2008) when he examined the participation of children in Mental Health Care and concluded that participation could helps to improve the quality and the effectiveness of Mental Health Care, because children can express their views about their own concerns and priorities. In addition, children become more confident with better knowledge about the healthcare system, so they do not only increase their relationships with clinicians, but also they can influence and control their experiences. Finally, children feel more respected by others, more confident about themselves, which have as results to improve their communications skills and their self-esteem.

Almost the same positive outcomes applied for youth who have experienced violence. If these children have the opportunities to participate in decision-making they become more confident with higher self-esteem. Furthermore, they acquire skills in order to speak, understand and finally overcome that traumatic experience. All these skills which will be made them stronger and more capable to protect themselves. Finally, participation of these children in decision-making about themselves can reduce the sense of social exclusion and on the other hand to increase their social and political capability (Horwath et al., 2012).

The outcomes of participation in decision-making for disabled children, is still an underdevelopment area, because it is difficult to measure it in that sample. Although, some studies, such as by Franklin and Sloper (2009) and Cavet and Sloper (2009), noted that all these outcomes which are applied in children with typical development are applied also in children with developmental disabilities. More specific, both studies concluded that children who participate in decision-making feel more valued, that they were listened by other people, more confident about themselves and able to acquire new skills. Moreover, there were evidences that these children after participation were more vulnerable to pay attention to their adolescence and to their lifestyle, but less vulnerable to a potential abuse.

5 Barriers to participation of children

As mentioned above, many are the benefits for children through participation in decision-making. However, many are the barriers of adults to the implementation of the Article 12 and to the non-participation of children in participation process (Hinton, 2008; Lundy, 2007; Matthews et al. 1999; Saaltink, Mackinnon, Owen, & Tardif-Williams, 2012; Tisdall, Davis & Gallagher, 2008). According to Lundy (2007) adults can separated into three groups; first them who are scepticism about children's abilities, so they believe that children lack capacity to have real participation in decision-making process, *with this conclusion came up the research by* Saaltink, Mackinnon, Owen and Tardif-Williams, (2012) who examined family members' perceptions for the participation of young people with intellectual disabilities (ID), in the family context. The second group are they who believe that if as adults give more control to children; that will decrease their power/authority and will destabilise the school environment (when we refer in decision-making into school environments). Finally, adults have the opinion that the more the participation of children the more will must be the efforts.

Before Lundy at 2007, Matthews et al. 1999 also noted three factors, which contribute to the non-participation of children in decision-making in relation to their representation in society, which are relative to the decision-making about themselves and similar to these factors of Lundy. First of all, adults thinking the appropriate situation and the appropriate level of involvement, second they believe that children are not capable and finally they are uncertain about the form of the participation and the results of a process like this.

Two more studies by Hinton (2008) and by Tisdall et al. (2008) highlighted the problem of participation process and noted some barriers which blocking that process. Their studies were broader and they examined the participation of children with typical development in organisation. However, some of these barriers could applied also in children with developmental disabilities and in decision-making process. These barriers are, first of all the existing rules and procedures which are crystallized in

adults minds and reinforce the existing procedures, according with lack of training. The attitudes towards the participation process of adults are more or less negative, according with no critical internal reflection. Finally, the bureaucracy along with the dominant/inaccessible language and staff's cynicism are also important obstacles (Coyne, & Gallagher, 2011; Kramer, Olsen, Mermelstein, Balcells, & Liljenquist, 2012).

6 Conclusion

To conclude, participation is a term which can take many forms and could explained with different ways. Although, as becoming clear from all the previous studies; participation of children with typical development in decision-making, is a process, which has many benefits for children's psychological and social development. However, adults and our modern societies blocked these processes with many barriers. Finally, it is really interesting to examine, if all these happened also in children with developmental disabilities, where the existing literature review on that topic is limited and if the benefits of participation in decision-making could also applied in these children along with techniques of empowerment.

References

Barker, J. (1999). Street-level democracy. West Hartford, CT: Kumarian Press.

Benard, B. (1993). Fostering resiliency in kids. *Educational Leadership*, *51*(3), 44–48. Boyden , J., Ennew , J. (1997). Children in Focus: A Manual for Participatory Research with Children.

Cavet, J., & Sloper, P. (2004). The participation of children and young people in decisions about UK service development. *Child: care, health and development, 30*(6), 613–21.

Coster, W., & Khetani, M. A. (2008). Measuring participation of children with disabilities: issues and challenges. *Disability and rehabilitation*, *30*(8), 639–48.

Coyne, I., & Gallagher, P. (2011). Participation in communication and decisionmaking: children and young people's experiences in a hospital setting. *Journal of clinical nursing*, 20, 2334–2343.

Day, C. (2008). Children's and Young People's Involvement and Participation in Mental Health Care. *Child and Adolescent Mental Health*, *13*(1), 2–8.

McNeish, D. (1999). Promoting participation for children and young people: some key questions for. *Journal of social work practice*, *13*(2), 191–203.

Franklin, A., & Sloper, P. (2009). Supporting the Participation of Disabled Children and Young People in Decision-making. *Children & Society*, 23(1), 3–15.

Gray, A. (2002). Increasing the participation of children, young people and young adults in decision making: A literature review.

Hammel, J., Magasi, S., Heinemann, A., Whiteneck, G., Bogner, J., & Rodriguez, E. (2008). What does participation mean? An insider perspective from people with disabilities. *Disability and rehabilitation*, *30*(19), 1445–60.

Hart, R. (1992). Children participation, from tokenism to citizenship. *Innocenti Essay.* UNICEF, 4.

Hart, R. A. (1999). Children's participation. New York: UNICEF.

Head, B. W. (2011). Why not ask them? Mapping and promoting youth participation. *Children and Youth Services Review*, *33*(4), 541–547.

Hinton, R. (2008). Children's Participation and Good Governance: Limitations of the Theoretical Literature. *The International Journal of Children's Rights*, *16*(3), 285–300.

Horelli, L. (1998). Creating Child-Friendly Environments: Case Studies on Children's Participation in Three European Countries. *Childhood*, *5*(2), 225–239.

Horwath, J., Kalyva, E., & Spyru, S. (2012). "I want my experiences to make a difference" promoting participation in policy-making and service development by young people who have experienced violence. *Children and Youth Services Review*, 34(1), 155–162.

Kramer, J. M., Olsen, S., Mermelstein, M., Balcells, A., & Liljenquist, K. (2012). Youth with disabilities' perspectives of the environment and participation: a qualitative meta-synthesis. *Child: care, health and development*, *38*(6), 763–77.

Kirby, P., & Bryson, S. (2002). *MEASURING THE MAGIC? Evaluating and researching public decision making MEASURING THE MAGIC? Evaluating and researching young people's participation in public decision making.*

Lansdown, G. (2001). Promoting children's participation in democratic decisionmaking. Florence: UNICEF Innocenti Research Centre.

Lundy, L. (2007). "Voice" is not enough: conceptualizing Article 12 of the United Nations Convention on the Rights of the Child. *British Educational Research Journal*, *33*(6), 927–942.

Mannion, G. (2003). Children's participation in school grounds developments: creating a place for education that promotes children's social inclusion. *International Journal of Inclusive Education*, 7(2), 175–192.

Matthews, H., Limb, M., & Taylor, M. (1999). Young people's participation and representation in society. *Geoforum*, *30*, 135–144.

Moses, S. (2008). Children and Participation in South Africa: An Overview. *The International Journal of Children's Rights*, *16*(3), 327–342.

Rosenberg, L., Jarus, T., Bart, O., & Ratzon, N. Z. (2011). Can personal and environmental factors explain dimensions of child participation? *Child: care, health and development*, *37*(2), 266–75.

Saaltink, R., Mackinnon, G., Owen, F., & Tardif-Williams, C. (2012). Protection, participation and protection through participation: young people with intellectual

disabilities and decision making in the family context. Journal of intellectual disability research 56(11), 1076–86.

Sheet of the Governments Journal (FEK), N 2101 192/2-12-1992.

Shier, H. "Pathways to Participation: Openings, Opportunities and Obligations". Children and Society 2001, (15), 107-117.

Sinclair, R. (2004). Participation in practice: making it meaningful, effective and sustainable. *Children and Society*, 18, pp. 106–118.

Tisdall, K., Davis, J., & Gallagher, M. (2008). Reflecting on Children and Young People's Participation in the UK. *The International Journal of Children's Rights*, *16*(3), 343–354.

Treseder, P. (1997). *Empowering children and young people: training manual*. Save the Children and Children's Rights Office: London.

U.N. Convention on the Rights of the Child. (1989). UN General Assembly Document A/RES/44/25.

World Health Organization (2001) International Classification of Functioning, Disability and Health (ICF). World Health Organization, Geneva, Switzerland.

The Role of Optimism Bias in Greek Young Drivers' Propensity to Commit Road Violations

Triantafylli Danai1

¹ South-East European Research Centre (SEERC), datriantafylli@seerc.org

Abstract. Road accidents are a major cause of death and injury worldwide and younger drivers are generally over-represented in accident statistics. Previous findings indicate younger drivers tend to drive in riskier manners. Also, previous research has showed that drivers overall tend to overestimate their driving skills and safety and there is evidence that this is more so for younger drivers. The present study measured 108 Greek drivers' optimism bias, as well as their willingness to violate road safety rules. Correlational results showed that younger age is associated with increased willingness to violate. Regression analysis revealed that age is a significant predictor of violation willingness, but when optimism bias measures are included, the effect of age is no longer significant. These findings indicate that optimism bias may be behind the overrepresentation of young drivers in road accidents. Implications and suggestions for future research are discussed.

Keywords: young drivers, optimism bias, violation willingness

1 Introduction

Traffic accidents are a major public health hazard universally, claiming, without any warning signs or symptoms, 1.2 million lives per year worldwide (World Health Organization, WHO, 2004). In the European region alone, road accidents claim a total of 130.000 lives annually and cause over 2.4 million injuries (WHO, 2009). Greece is consistently among the EU countries with the most road accident fatalities, with approximately 15 deaths per 100.000 people, while the EU average is at 13.4 fatalities per 100.000 people (WHO, 2009). In 2012, speeding, red light violations, illegal overtaking and other examples of risky and reckless driving behaviour caused 33% of all road accident fatalities in Greece; i.e. given a more sensible driving behaviour 322 lives would have been spared (Hellenic Police, 2012).

1.1 Age, crash risk, and optimism bias

Worldwide the age of the driver has been found to be one of the most important factors with regard to accident involvement. Specifically, data from USA reveal that crash rates (both minor and more serious, leading to injuries and fatalities) for drivers aged from 16 to 24, are the highest of all age groups (National Highway Traffic Safety Administration, NHTSA, 2001). Data from the UK show that male drivers aged from 17 to 20 are involved in 334 injury accidents per 100 million km driven more than the average of all male drivers, and although the difference declines for the age group 20-24, the accident occurrences for this age group are still 70% above the all male average (Forshyth, 1992, as cited in Clarke, Ward, & Truman, 2005). Additionally, according to the Australian Bureau of Statistics (2008) road accidents are the leading cause of death among people younger that 25 years of age. Finally, a crosscultural study of driving behaviour, which included Greece, Iran, Turkey, Finland, the Netherlands and the UK, showed that for all these countries age was significantly correlated with safe driving skills (Özkan, Lajunen, Chliaoutakis, Parker, & Summala, 2006). As such it appears that "there is something about being young that puts young drivers at high risk for crash involvement" (Arnett, 2002, p. ii17).

Research on the topic of the over-representation of young drivers in accident involvement has shown that factors such as reduced experience, driving conditions, contribution of alcohol consumption, as well as the condition of the car, cannot sufficiently explain this over-representation (Hodgson, Bragg, & Finn, 1981, as cited in Matthews & Moran, 1986). A more recent study showed that although the behaviours that contribute to accident involvement of young drivers are practically identical with those of older age groups, drivers of this age group exhibit an increased tendency to overlook routine safety practices while driving and to ignore or underestimate the resulting danger (McKnight & McKnight, 2003). Surprisingly, evidence suggests that this tendency cannot be attributed to lack of awareness with regard to risk, since young drivers report that they know they belong to a high-risk age group and that they understand the relevant risks, while at the same time report that they are not personally susceptible (Guerin, 1994; Harré, Foster, & O'Neill, 2005).

This relinquishing of personal relevance, this tendency to overestimate personal ability and underestimate personal risk is known as self-enhancement bias, or optimism bias (Weinstein, 1980), and it is not exclusively met in young drivers. Relevant research has shown that drivers of all age groups tend to rate themselves as significantly safer, slower, and more skillful that the average driver (e.g. Gosselin, Gagnon, Stinchcombe, & Joanisse, 2010; Harré & Sibley, 2007; Horswill, Waylen, & Tofield, 2004). Nevertheless, this phenomenon is more pronounced in younger drivers (e.g. Tränkle, Gelau, & Metker, 1990), and it is possible that the overrepresentation of young drivers in road accidents can be explained, at least partially, by their increased engagement in self-enhancement bias.

1.2 The present study

Taking as a starting point the fact that beliefs resulting from engagement in optimism bias have been associated with greater risk of involvement in accidents (Clarke, Ward, & Truman, 2005; Harré, 2000), the present study aims to explore the relationship between the age of the driver, their engagement in optimism bias with regard to driving skills, safety, and crash involvement risk, and their driving behaviour. Since it was not possible to measure actual driving behavior in this study, a way to measure future behavior needed to be employed. For that purpose, the author looked to influential behavioural theories, such as the Theory of Reasoned Action (Fishbein & Ajzen, 1980) and the Theory of Planned Behaviour (Ajzen, 1985). According to those theories the key predicting component of behavior is the concept of behavioural intention. However, behavioural intention is defined as the decision to engage or not in a certain behavior, reached after of a goal-oriented process, where available behavioural options along with their possible results and consequences are evaluated (Gibbons, Gerrard, Blanton, & Russel, 1998) and as such it is based on an assumption that a certain level of rational forethought is involved in the process. With regard to behaviours related to health risk it has been argued (and supported by research findings) that young people (adolescents and young adults alike) do not behave in a rational or intentional manner (Prototype/willingness model; Gibbons & Gerrard, 1995, 1997; Gibbons, Gerrard, Ouelette, & Burzette, 1998); rather, they tend to react to the available risk-conducive conditions.

Consequently, in order to circumvent the assumption of rational forethought that is embedded in the concept of behavioural intention, in the present study the researcher chose to employ the concept of behavioural willingness (BW). BW has "been developed to capture situations in which people typically plan or intent to behave in one way, but may be prepared to behave differently should the circumstances permit or sanction" (Rowe, Andrews, & Harris, 2013, p.3). In particular, this research project utilized the Violation Willingness Scale (VWS; Rowe, Andrews, & Harris, 2013) in order to measure how willing the participants would be to violated certain road safety rules, rather than to measure how intent they are on violating them.

Research questions and hypotheses:

RQ1: Will optimism bias reduce as a function of increasing drivers' age?

H₁: Optimism bias scores will be negatively correlated with the participants' age.

RQ2: Will the propensity to violate be a function of both drivers' age and their engagement in optimism bias?

H₂: Age and optimism bias scores will predict willingness to violate scores.

2 Method

2.1 Participants

Participants were recruited from the staff and the student body of City College, Affiliated Institution of the University of Sheffield, and of the South-East European Research Centre (SEERC). Also, random recruitment of eligible participants on the streets of Thessaloniki was be performed in order to ensure a sufficient number of participants is available. The participation criteria were to hold a car driving licence and they drive a car at least 2 days per week. A total of 108 people participated in the study, with mean age 33.7 (SD= 11.95, range = 19-62), 83 of them were male and 25 were female. The male-female ratio in this study closely resemble the gender ratio of licensed drivers in Greece (Focus Bari, 2010). The study was approved by the Department of Psychology, University of Sheffield Ethics committee, and all participants signed an informed consent prior to participation.

2.2 Measures

All measures used in the study were translated from English to Greek, and then back-translated to ensure accuracy of translation. The Greek versions were used for data collection.

Demographics. Participants were asked to provide the following information about themselves: age, sex, years driving, and frequency of driving.

Optimism Bias. For the measurement of drivers' optimism bias some of the items used for the same purpose in Harre and Sibley (2007) were employed. The question "Do you think you are more or less skilled as a driver than other people of your age and gender?" was used to measure optimism bias for driving skill. The same question substituting "safe" for "skilled" was used for optimism bias for driving safetyFor optimism bias with regard to crash risk, the question "Do you think you are more or less likely to be involved in an accident while driving than other people of your age and gender?". All the questions were answered on a scale ranging from 1 - much less than others, to 7 - much more than others, with the question referring to crash risk being reversely scored.

Willingness to engage in road violations. The Violation Willingness Scale (VWS; Rowe, Andrews, & Harris, 2013) was used to measure willingness to engage in road violations. The scale consists of 8 hypothetical scenarios, each of them corresponding to the violations measured by the Driving Behaviour Questionnaire (Reason, Manstead, Stradling, Baxter, & Campell, 1990). The VWS provides scenarios such as "You are running late for an appointment" and asks respondents to rate on a 7-point scale ranging from 1 - not likely at all, to 7 - very likely. how likely it would be to react in various manners such as "Break the speed limits while driving in a town", "Break the speed limits while driving on country roads", "Break the speed limits while driving on a motorway".

3 Results

3.1 Violation Willingness (VW)

A Chronbach's Alpha analysis was employed to investigated whether a single VW score could be reliably constructed. The analysis showed that the items of the VWS formed a scale with high reliability (alpha = .826), and thus the single variable VW was created using the mean of the answers to the VWS.

3.2 Relationship Between Age, Optimism Bias and VW

The correlation between age and the three different optimism bias measures was examined, as well as the correlation between age and VW scores. There was no significant correlation between age and any of the three optimism bias measures (Table 1). There was however a significant negative correlation between age and VW score (r = -.257, p < .01).

Table 1.: (Correlations	between	age and	optimism	bias measures
-------------	---------------------	---------	---------	----------	---------------

Measure	Bias_Skill	Bias_Safe	Bias_Accident
Age	<i>r</i> =044, <i>p</i> = .651	r = .152, p = .116	r = .063, p = .515

3.3 Age and Optimism Bias measures as Predictors of VW scores

Using the enter method, a regression analysis was performed with VW scores as the dependent variables, and age, gender, years of driving experience and the three optimism bias measures as predictor variables. The first block of predictor variables to be entered were age, gender and years of driving experience. This first model was significant: F(3,103) = 3.180, p < .05, and it explained 5.8% of the variance (Adjusted $R^2 = .058$), with age emerging as the only significant predictor (Table 2.1).

Variable	В	SE B	β
Age	061	.03	724*
Gender	.045	.25	.019
Years Driving	.044	.03	.488

Table 2.1.: Regression analysis Model 1

**p* < .05

The second block of variables to be added to the model was comprised of the three optimism bias variables: bias_skill, bias_safe, and bias_accident. The new model that emerged was also significant: F(6,100) = 4.965, p < .001, and it explained 18.3% of the variance (Adjusted $R^2 = .183$). However, with the addition of the new variables age was no longer a significant predictor and the measures of bias regarding skill (Bias_Skill) and regarding safety (Bias_Safe) emerged as significant (Table 2.2).

Variable	В	SE B	β
Age	051	.029	601
Gender	017	.236	007
Years_Driving	.039	.032	.432
Bias_Skill	.232	.086	.280*
Bias_Safe	316	.077	414**
Bias_Accident	.050	.070	.065

 Table 2.2: Regression analysis Model 2

**p* < .05

***p* < .01

4 Discussion

The present study aimed at investigating whether age plays a significant role in Greek drivers' engagement in optimism bias regarding their skills, safety, and crash risk, as well as in drivers' willingness to violate legal and conventional safety rules. The results of the analyses showed that there was a negative relationship between age and violation willingness scores, which means that as drivers' age increases their willingness to violate the aforementioned safety rules decreases. Additionally, although there was no direct relationship between age and engagement in self-enhancement biases found, further analysis revealed that when self-enhancement biases regarding skills and safety (and not regarding crash risk) are taken into account the effect of age on violation willingness is no longer significant, partially confirming the second hypothesis of the study. This finding can be interpreted as that the effect that the driver's age has on his/her willingness to violate safety standards can be partially explained by his/her engagement in optimism biases.

The first finding of the study is not surprising, since younger age has consistently been found to be associated with decreased precautionary driving (e.g. Özkan et al., 2006); it does however lend additional support that this relationship holds true for Greek drivers. The second finding, viewed simply as the prediction of violation will-ingness by drivers' optimism biases was also in line with previous research findings (e.g. Clarke, Ward, & Truman, 2005; Harré, 2000). However, when taking into account the fact that the addition of the optimism bias measures to the predictive model

for violation willingness the effect of age was rendered obsolete, the specific finding assumes greater importance.

What is of particular interest in these findings is the negative contribution of the self-enhancement measure regarding safety to the final predictive model. It appears that participants are aware that their increased willingness to violate road safety rules results in decreased safety while driving; yet, they still report willing to violate. Nevertheless, taken together with the positive contribution of bias regarding driving skill, it can be argued that participants rely on what they believe to be increased driving skill in order to self-justify their willingness to engage in driving behaviours that they know to be risky and reckless.

This is the first study, to the author's knowledge, to empirically demonstrate that it is possible that the mechanism responsible (at least partially) for young drivers' unsafe behaviour on the roads is due to their biased beliefs about their driving abilities, but not about their driving safety or about their personal risk to be involved in an accident. However, the sample of this study was rather small, and as such further studies, with larger samples, are needed in order to examine the replicability of the finding. Assuming that the results of this study will be supported by future study findings, then it becomes clear that in order to make the roads safer and to also reduce young people's injuries and deaths, methods of minimizing or extinguishing these biased beliefs need to be explored.

Additionally, this was the first study to employ the concept of violation willingness as an antecedent to actual driving behaviour. The fact that increased propensity towards violations was associated with younger age closely resembles the welldocumented propensity of younger drives to drive less safely and as such provides support to the argument that violation willingness can be used as an antecedent to driving behaviour. It would however be useful if future studies examined it in relation to both intentions to violate and actual driving behaviour.

4.1 Conclusions

It appears that young drivers' optimism bias regarding their driving skills in particular is a contributory reason why this age group is over-represented in road accident statistics universally. This is important because, if replicated, it can provide a starting point towards reducing those grim statistics, and new interventions aimed at reducing self-enhancement biases can be designed to address the problem. Additionally, the findings provide some initial support for the use of violation willingness in future studies of driving behaviour, where measurement of actual driving is not possible and researchers rely solely on self-report measures.

References

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior, in Kuhl. J. & Beckman, J. (Eds.), *Action control: From cognition to behavior*, Springer-Verlag, Berlin, pp. 11-39

Arnett, J. J. (2002) Developmental sources of crash risk in young drivers. *Injury Prevention*, 8 (Suppl II), ii17-ii23

Australian Bureau of Statistics, 2008. Australian social trends, 2008: Risk taking by young people. Retrieved May 22, 2013 from http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4102.0Chapter5002008

Clarke, D. D., Ward, P., & Truman, W. (2005) Voluntary risk taking and skill deficits in young driver accidents in the UK. *Accident Analysis and Prevention*, 37 (3), 523-529

Fishbein, M., & Ajzen, I. (1980). Predicting and understanding con- sumer behavior: Attitude-behavior correspondence, in Ajzen, I. & Fishbein, M. (Eds.), *Understanding attitudes and predicting social behavior*, Prentice-Hall, Englewood Cliffs, pp. 148-172

Gibbons, F. X., & Gerrard, M. (1995) Predicting young adults' health- risk behavior. *Journal of Personality and Social Psychology*, 69, 505-517

Gibbons, F. X., & Gerrard, M. (1997) Health images and their effects on health behavior. In B. P. Buunk & F. X. Gibbons (Eds.), *Health, coping, and well-being: Perspectives from social comparison theory*, Mahwah, NJ: Erlbaum, pp. 63-94

Gibbons, F. X., Gerrard, M., Blanton, H., & Russel, D.W. (1998) Reasoned action and social reaction: willingness and intention as independent predictors of health risk. *Journal of Personality and Social Psychology*, 74 (5), 1164-1180

Gibbons, F. X., Gerrard, M., Ouelette, J., & Burzette, B. (1998). Cognitive antecedents to adolescent health risk: discriminating between behavioral intention and behavioral willingness. *Psychology and Health*, 13, 319-340

Gosselin, D., Gagnon, S., Stinchcombe, A., & Joanisse, M. (2010) Comparative optimism among drivers: an intergenerational portrait. *Accident Analysis and Prevention*, 42(2), 734-740 Guerin, B. (1994) What do people think about the risks of driving? Implications for traffic safety interventions. *Journal of Applied Social Psychology*, 24 (11), 994-1021

Harré N., Foster, S., & O'Neil, M. (2005). Self-enhancement, crash-risk optimism and the impact of safety advertisements on young drivers. *British Journal of Psychology*, 96 (2), 215 – 230

Harré, N. (2000) Risk evaluation, driving and adolescents: a typology. *Developmental Review*, 20, 206-226

Harré, N., & Sibley, C. G. (2007). Explicit and implicit self-enhancement biases in drivers and their relationship to driving violations and crash-risk optimism. *Accident Analysis and Prevention*, 39 (6), 1155–1161

Hellenic Police (2012) Statistical data of traffic police for the year 2012. Retrieved May 23, 2013, from http://www.astynomia.gr/index.php?option=ozo_content&perform=view&id=10533 &Itemid=86&lang=

Horswill, M. S., Waylen, A. E., & Tofield, M. I. (2004) Drivers' ratings of different components of their own driving skill: A greater illusion of superiority for skills that relate to accident involvement. *Journal of Applied Social Psychology*, 34(1), 117-195

Matthews, M. L., & Moran, A. R. (1986) Age differences in male drivers' perception of accident risk: the role of perceived driving ability. *Accident Analysis and Prevention*, 18 (4), 299-313

McKnight, J.A., & McKnight, S. A. (2003) Young novice drivers: careless or clueless? *Accident Analysis and Prevention*, 35, 921-925

National Highway Traffic Safety Administration (2001) *Traffic Safety Facts, 2000,* NHTSA, Washington, DC

Rowe, R., Andrews, E., & Harris, P. (2013) Measuring risky-driving propensity in pre-drivers: the Violation Willingness Scale. *Transport Research Part F*, 19, 1-10.

Tränkle, U., Gelau, C., & Metker, T. (1990) Risk perception and age specific accidents of young drivers. *Accident Analysis and Prevention*, 22, 119-125

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, 39 (5), 806-820

World Health Organization (2004) World Report on Road Traffic Injury Prevention, WHO, Geneva

World Health Organization (2009) European Status Report on Road Safety: Towards Safer Roads and Healthier Transport Choices, WHO Regional Office for Europe, Copenhagen

Özkan, T., Lajunen, T., Chliaoutakis, J. E., Parker, D., & Summala, H. (2006) Crosscultural differences in driving skills: a comparison of six countries, *Accident Analysis and Prevention*, 38, 1011-1018

Author Index

Abruev, Akmal, 19 Aggelopoulos, Stamatis, 19 Anastasiadou, Dimitra, 377 Aranitasi, Marin, 224, 275 Avramova, Andrijana Popovska, 354 Axtell, Carolyn, 399

Basholli, Adelina, 233 Bexheti, Lejla Abazi, 284 Bratanis, Konstantinos, 301

Cini, Mirela, 50 Coyne, Iain, 399 Cucllari, Frederik, 50 Cvetanovski, Ilija, 32

Dautov, Rustem, 247 Dhimitri, Eva, 50 Drabova, Jana, 386 Draganac, Dragana, 64 Drencheva, Andreana, 78

Edmond, Jajaga, 284 Ejodame, Ehimen, 259

Farley, Sam, 399 Frasheri, Neki, 331

Gajdokova, Tereza, 386 Gonidis, Fotis, 270 Gourgoura, Esida Gila, 90 Graell, Montserrat, 377 Grigoriadou, Thaleia, 404 Gurba, Krystian, 107 Gusev, Marjan, 313

Iakovidou, Olga, 188 Ibro, Marsida, 224, 275 Ismaili, Flaka, 113

Kalyva, Efrosini, 506 Kapoulas, Alexandros, 32 Kiroski, Kiril, 313 Knepper, Paul, 421 Kostoska, Magdalena, 313 Koturovic, Darja, 413 Kourtesis, Dimitrios, 301 Kushe, Renalda, 321 Kyprianou, Voula, 421 Kyriakidis, Alexandros, 433 Kyrkou, Artemis, 467

Lagkas, Thomas, 233 Limani, Fidan, 294 Lule, Ahmedi, 284

Maguire, Stuart, 259 Melonashi,Erika, 495 Milosevic, Sinisa, 123 Mocanu, Mihai, 365 Mocka, Blerta, 275 Murtazaev, Olim, 19

Naaranoja, Marja, 172 Nikolaidou, Eftychia, 90

Paraskakis, Iraklis, 247, 270 Pavlidis, George, 476 Popa, Radu Teodoru, 365 Proedrou, Filippos, 421

Qehaja-Osmani, Fatbardha, 489

Radevski, Vladimir, 294 Rerichova, Vlasta, 386 Ristov, Sasko, 313 Rukiqi, Linda Tahiri , 139

Sepulveda, Ana Rosa, 377 Serafini, Giovanni Oscar, 145 Shehu, Adrian, 321 Shehu, Bojken, 224 Sheme, Enida, 331 Shkullaku, Rudina, 495 Siakas, Kerstin, 172 Simons, Anthony J. H., 270 Sprigg, Christine, 399 Stannet, Mike, 247 Stojkoska, Biljana, 340, 354

Tasios, Theodosios, 506 Triantafylli, Danai, 516

Vavili, Fani, 467 Vivas, Ana, 476 Vlachakis, Sotirios, 172 Vlachou, Harissia, 188

8th Annual South-East European Doctoral Student Conference

SEERC is a research centre of the International Faculty of the University of Sheffield, CITY College.

Mission

SEERC's mission is to support the development of South-East Europe by conducting pure and applied research in and for the region. To accomplish this, SEERC employs the existing research capacities of the University of Sheffield and CITY College- International Faculty of The University of Sheffield by facilitating collaborations between their research staff and by developing multi-disciplinary networks of researchers from across South-East Europe.

SEERC contributes to knowledge capacity building:

- By promoting a dialogue among SEE countries related to the development of the SEE region.
- By providing a forum for researchers in the SEE region.
- By disseminating research findings to academics, professionals, NGOs, and policy-makers through conferences, advising, published research papers, monographs, and working papers.
- By developing multi-disciplinary research perspectives through workshops and seminars.
- By promoting the integration of researchers from universities and research centres in the SEE region into networks of excellence.



