Volume 4, Number 3, 2017

REGION

The Journal of ERSA
Powered by WU

Table of Contents

– continued on next page –

Funded by

Table of Contents – continued

REGION, The journal of ERSA / Powered by WU
ISSN: 2409-5370

ERSA: http://www.ersa.org
WU: http://www.wu.ac.at

Editorials

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 3, 2017, E1–E5
DOI: 10.18335/region.v4i3.201

journal homepage: region.ersa.org

Special Issues at REGION

Vicente Royuela[1]

[1] University of Barcelona, Barcelona, Spain (email: vroyuela@ub.edu)

Received: 7 August 2017/Accepted: 7 August 2017

Starting a new academic journal is not an easy task. The production function of researchers has as inputs many hours of work, with the output being a small countable number of publishable pieces. Consequently, every researcher must evaluate carefully where to submit their output for publication. From time to time we see calls for papers or receive invitations to contribute to Special Issues. Is that a good outlet for our papers? It is likely our papers will capture a more specialized audience if published in a Special Issue and consequently, one can expect these papers to be more frequently cited. Nevertheless, as contributions to Special Issues are usually published all together, faster papers are held up until the last paper of the volume is accepted for publication.

Every scientific discipline surely has its own characteristics and probably this way of wrapping papers into special issues can be a good idea or a bad strategy depending on the domain. What is the situation in Regional Science? Are Special Issues a good way to advertise our research? To answer this question, I have collected information about six journals[1] linked to the Regional Science Association International (Papers in Regional Science and Regional Science Policy and Practice) or to ERSA sections (Investigaciones Regionales, Jahrbuch für Regionalwissenschaft, Scienze Regionali and Spatial Economic Analysis) over the last ten years (2007-2016). All six journals are indexed in Scopus and one can trace the citation record of every paper. I have discarded editorials, notes and articles in press. Table 1 displays the descriptive statistics of the citations of the 1,114 considered papers. As can be expected, Papers in Regional Science, the flagship journal of RSAI, is the publication with most citations and published papers. The modal citation is zero, although younger publications (such as Regional Science Policy and Practice) tend to collect less citations.

Every journal except Regional Science Policy and Practice had at least one Special Issue in the analysed period. Figure 1 shows that papers published in Special Issues accounted for 29% of papers published in 2014, while in 2016 they represented a mere 10%, confirming an absolute decline since the maximum number in 2014.

To find out whether contributions to Special Issues attract more citations, I have regressed the citation against a dummy for every journal, the publication year and a dummy for the label Special Issue. Table 2 summarises the results. On average, papers published in Special Issues are significantly more cited than regular papers. This is particularly true for articles published in Papers in Regional Science, while we can only find parameters that are not significant for the other journals. Of course, one could account for several other dimensions, such as the number of authors, the length of the papers, and even the quality of authors (for instance adjusting by the h-index of every

[1] I have collected the full period for Investigaciones Regionales, Jahrbuch für Regionalwissenschaft, Papers in Regional Science and Spatial Economic Analysis. For Scienze Regionali we consider papers published since 2012 and for Regional Science Policy and Practice since 2015.

Table 1: Citation records in Regional Science journals

Cites per paper	Investigaciones Regionales	Jahrbuch für Regional-wissenschaft	Papers in Regional Science	Regional Science Policy and Practice	Scienze Regionali	Spatial Economic Analysis	Total
0	127 (50,6%)	36 (40,4%)	80 (18,1%)	20 (76,9%)	60 (58,3%)	11 (5,4%)	334 (30%)
1	50 (19,9%)	12 (13,5%)	60 (13,5%)	4 (15,4%)	21 (20,4%)	39 (19,3%)	186 (17%)
2-5	64 (25,5%)	31 (34,8%)	124 (28%)	0 (0%)	13 (12,6%)	79 (39,1%)	311 (28%)
6-10	7 (2,8%)	9 (10,1%)	72 (16,3%)	2 (7,7%)	1 (1%)	33 (16,3%)	124 (11%)
11-20	3 (1,2%)	1 (1,1%)	60 (13,5%)	0 (0%)	5 (4,9%)	21 (10,4%)	90 (8%)
21-50	0 (0%)	0 (0%)	32 (7,2%)	0 (0%)	3 (2,9%)	15 (7,4%)	50 (5%)
51-100	0 (0%)	0 (0%)	11 (2,5%)	0 (0%)	0 (0%)	2 (1%)	13 (1%)
>100	0 (0%)	0 (0%)	4 (0,9%)	0 (0%)	0 (0%)	2 (1%)	6 (0,5%)
Total Papers	251 (100%)	89 (100%)	443 (100%)	26 (100%)	103 (100%)	202 (100%)	1,114 (100%)
Total Citations	338	199	4,355	19	233	1,776	6,920

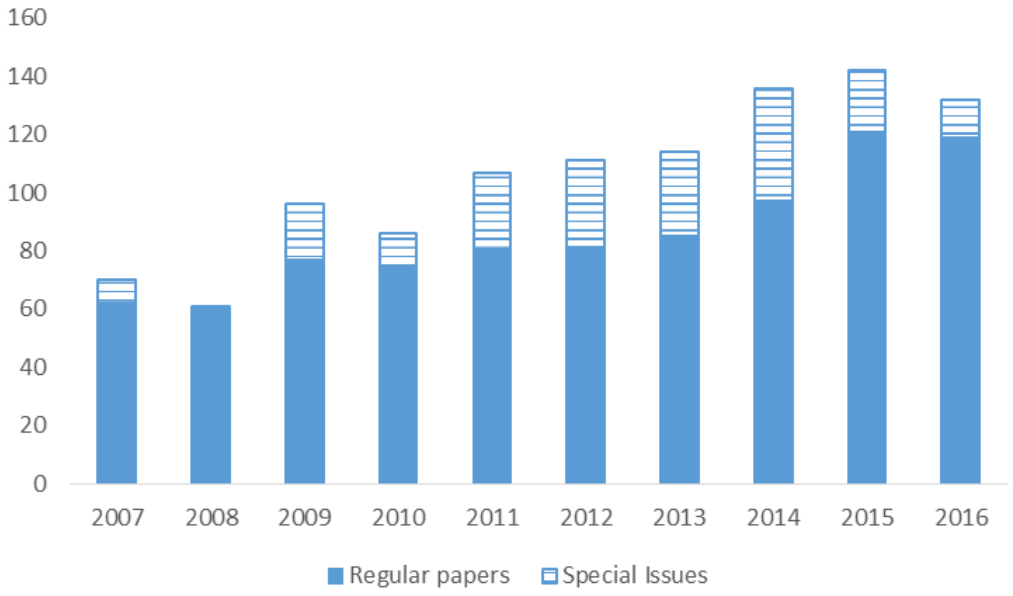


Figure 1: Papers published in regular volumes versus Special Issues

author). Still, this brief exercise indicates that Special Issues do matter and on average have a positive impact on the number of citations.

In traditional journals, this positive effect of Special Issues comes at a price. With strict limits for the numbers of pages or issues per year, allocating space to a Special Issue crowds out regular contributions. Successfully refereed articles are bypassed by the Special Issue and must wait even longer in the journal's backlog for proper publication.

REGION is particularly active in the promotion of Special Issues, since the electronic format allows us to reap the benefits while avoiding the drawbacks. We expect Special Issues to attract high quality papers and to subsequently further improve the standards of the journal. Of course, we are too young to be included in the more important indexing databases, and consequently it is not easy to capture the impact of our editorial policy on citations. With our editorial policy, we try to maintain the advantages of Special Issues while avoiding the disadvantages. Since REGION publishes only electronically, there are no page or issue constraints in place and therefore, no regular papers are crowded out. Moreover, we apply the policy that once a paper – regular or special issue – is accepted, it gets published immediately. Consequently, faster authors are not held back by slower the ones. Once the last Special Issue paper is published, we bundle them and a volume editorial together into a separate volume advertising the published papers, which are

referenced to their original issue: "all contributions to this special issue have already been published in earlier issues of REGION, for the sake of immediate exposure of the content." We believe that authors, special issue editors as well as the journal benefit from this policy.

Right now REGION has published its first Special Issue as Volume 4 number 2 on "Well-being in cities and regions: measurement, analysis and policy practices", edited by Paolo Veneri (OECD) and Arjen J.E. Edzes (University of Groningen) and we have also published papers from several other Special Issues and some new calls are open. We welcome new proposals!

Table 2: Citation patterns by journal and publication year

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	All	Investigaciones Regionales	Jahrbuch für Regional-wissenschaft	Papers in Regional Science	Regional Science Policy and Practice	Scienze Regionali	Spatial Economic Analysis
Special Issue		4.7829***	0.1977	-1.2381	8.7997***		-1.0603	4.46286
		(1.203)	(0.344)	(2.061)	(2.437)		(1.290)	(4.269)
Journal - Reference category: Investigaciones Regionales								
Jahrbuch für Regionalwissenschaft	0.0258	1.1081						
	(1.868)	(1.875)						
Papers in Regional Science	9.4484***	9.8664***						
	(1.200)	(1.196)						
Regional Science Policy and Practice	5.7092*	6.7039**						
	(3.340)	(3.325)						
Scienze Regionali	4.9786***	5.1728***						
	(1.808)	(1.796)						
Spatial Economic Analysis	8.0222***	8.5429***						
	(1.444)	(1.440)						
Year: reference category 2007								
2008	-5.2375**	-4.7192*	-0.6518	-0.6111	-11.3065**			1.9545
	(2.594)	(2.579)	(0.715)	(1.452)	(5.318)			(7.565)
2009	-9.4496***	-9.8094***	-1.1333*	-1.2540	-16.9445***			-9.3690
	(2.327)	(2.312)	(0.631)	(1.505)	(4.457)			(7.474)
2010	-3.4825	-3.5919	-1.6476**	-0.5556	-12.8016***			4.8045
	(2.386)	(2.370)	(0.729)	(1.408)	(4.445)			(7.250)
2011	-11.1299***	-11.7634***	-1.6722***	-2.8111**	-22.2874***			-6.9414
	(2.275)	(2.264)	(0.612)	(1.373)	(4.448)			(7.260)

Continued on next page

Table 2 – continued from previous page

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	All	Investigaciones Regionales	Jahrbuch für Regional-wissenschaft	Papers in Regional Science	Regional Science Policy and Practice	Scienze Regionali	Spatial Economic Analysis
2012	-12.3583***	-13.0731***	-0.5134	-1.3111	-22.7681***			-10.9830
	(2.275)	(2.266)	(0.653)	(1.373)	(4.542)			(7.403)
2013	-14.8453***	-15.4928***	-1.4003**	-2.8611*	-26.5940***		-1.6714	-14.9005**
	(2.259)	(2.249)	(0.664)	(1.452)	(4.385)		(1.938)	(7.275)
2014	-14.7971***	-15.5848***	-1.9430***	-1.5397	-29.6592***		4.6759***	-13.8461*
	(2.199)	(2.193)	(0.643)	(1.505)	(4.267)		(1.651)	(7.260)
2015	-17.2499***	-17.4027***	-1.7443***	-3.2778**	-29.0208***		-1.7717	-16.4567**
	(2.209)	(2.194)	(0.643)	(1.574)	(4.350)		(1.614)	(7.403)
2016	-18.1116***	-18.0576***	-2.5220***	-3.6111**	-31.3905***	-1.5833*	-1.8122	-15.6455**
	(2.215)	(2.199)	(0.632)	(1.452)	(4.267)	(0.869)	(1.917)	(7.250)
Constant	12.7788***	11.8381***	2.7143***	4.1111***	31.0208***	1.5833**	2.3380*	17.5455***
	(1.932)	(1.933)	(0.470)	(0.996)	(3.484)	(0.601)	(1.335)	(5.823)
Obs	1,055	1,055	243	85	414	23	103	187
R^2	0.171	0.183	0.100	0.155	0.219	0.137	0.211	0.126

Articles

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 3, 2017, 1–17
DOI: 10.18335/region.v4i3.142

journal homepage: region.ersa.org

Territory and Sustainable Tourism Development: A Space-Time Analysis on European Regions

João Romão[1] João Guerreiro[2] Paulo M. M. Rodrigues[3]

[1] University of Algarve, Faro, Portugal (email: jfromao@ualg.pt)
[2] University of Algarve, Faro, Portugal (email: jguerreiro@ualg.pt)
[3] Bank of Portugal and Universidade Nova de Lisboa, Lisbon, Portugal (email: pmrodrigues@bportugal.pt)

Received: 8 July 2016/Accepted: 26 April 2017

Abstract. In the long run, tourism competitiveness depends on the sustainable use of territorial assets: the differentiation of destinations depends on the integration of cultural and natural resources into the tourism supply, but also on their preservation over time. Using advanced spatial econometric techniques this work analyses the relationships between regional tourism competitiveness, the dynamics of tourism demand and investment, as well as the existence of natural resources and cultural assets in the European context. Despite the close relationship between tourism activities and the characteristics of the territories, the application of methods of spatial analysis in tourism studies is still scarce and the results of this work clearly show their potential for this field of research. Among the main findings of this paper, it was observed that natural resources do not have the expected positive impacts on regional tourism competitiveness and that European regions with more abundant natural resources are often developing unsustainable forms of mass tourism with low value added and little benefits for the host communities. The existence of spatial correlation effects suggests that positive spillovers arising from tourism dynamics in neighbourhood regions prevail over potential negative effects related to the competition between destinations. Policy and managerial implications of these results are discussed and further research questions are proposed.

Key words: Cultural heritage; Natural resources; Endogenous resources; Spatial autocorrelation; Spatial econometrics; Competitiveness; Sustainability

1 Introduction

As one of the fastest growing sectors in the contemporary global economy, tourism activities face unprecedented levels of competition. However, it is not only a competition between product and service providers, but also between destinations and, consequently, between territories. On the other hand, as tourism attractiveness often relies on territorial natural and cultural endowments, questions related to the sustainable use of these resources assume greater importance as tourism is achieving higher relevance at the international level. This is expressed by the number of tourists worldwide and also by their socio-economic and environmental impacts.

The purpose of this work is to provide a comprehensive analysis of the importance of local resources – both cultural and natural – for regional tourism competitiveness, measured in terms of the gross value added by the tourism sector, at the European level.

Our analysis is framed by the concepts of competitiveness and sustainability in tourism. We assume the definition of competitiveness proposed by Ritchie, Crouch (2003), which establishes a clear link to the socio-economic benefits for the local communities and the sustainable use of sensitive territorial resources. Section 2.1 offers a brief literature review for this conceptual framework focusing on the relations between competitiveness and sustainability in tourism.

As our study aims to identify the impacts of tourism on regional socio-economic dynamics, the territorial level of analysis is not exactly the destination, but rather the regional level (NUTS 2, according to the definitions of Eurostat, which is the territorial level typically used for the application of regional policies, while the NUTS 1 level corresponds to the major socio-economic regions and NUTS 3 to small regions). Although these regions normally include more than one tourism destination, they are institutionally relevant in order to address policy questions related to the integration of tourism dynamics into broader resource management or economic development policies. A synthesis of tourism competitiveness studies focused on this territorial level is presented in Section 2.1. In particular, Cucculelli, Goffi (2016) and Cuccia et al. (2016) have analysed similar problems related to natural and cultural resources.

Assuming tourism as a place oriented activity, where territories interact with each other, the possible existence of spatial effects among the regions under analysis is also of interest for the purposes of our study. Despite the close connection between tourism and the territorial characteristics, very few studies have applied spatial econometric methods to the field of tourism as described in Section 2.3. For the purpose of our study, the analysis of the relations between tourism dynamics and cultural assets in Italian regions proposed by Patuelli et al. (2013) or the study of the impacts of natural and cultural endowment on regional tourism demand at the European level developed by Romão (2015) are relevant examples.

Considering that competitiveness implies the creation of high value added and the sustainable use of natural and cultural resources, our paper applies spatial analysis techniques in order to identify different spatial patters in tourism dynamics (Section 3). This analysis is combined with an econometric overall explanation, with the estimation of a general trend within European regions (Section 4). As the work focuses on a large group of regions and considers a period of eight years, a spatial panel data model has been chosen in order to deal with cross-sectional and time series characteristics of the data, while allowing for the identification and quantification of potential spatial effects.

Although limited data availability constrains a comprehensive analysis of all questions related to regional tourism competitiveness, the available information and the methodologies applied in this study contribute to the understanding of the relations over time between attractive and sensitive territorial resources, tourism dynamics and tourism competitiveness at the European regional level. This work provides an innovative analysis of the spatial effects between tourism demand, tourism infrastructures, regional tourism competitiveness and the sustainable use of natural and cultural resources in European regions with relevant results arising from the exploratory spatial analysis and the regression model. A discussion of their policy implications will be presented in the concluding section and issues for further research will be highlighted.

2 Competitiveness, Sustainability and Spatial Econometrics in Tourism

2.1 Tourism competitiveness and sustainable use of resources

The concepts of competitiveness and sustainability emerged in the literature during the 1980s. Michael Porter's analysis (Porter 1985, 2003) regarding the achievement of a competitive advantage at the firm level had clear implications on economic policy formulations at the regional and national levels. On the other hand, the concept of sustainability was introduced after the publication of the document "Our Common Future" (World Commission on Environment and Development 1987), establishing the principles of sustainable development and taking into consideration its multiple dimensions (economic, social or environmental).

These concepts have been applied to tourism studies during the subsequent years and a good synthesis has been provided in a short note by Poon (1994), by applying some of the generic strategic formulations proposed by Porter and defining a strategy of cost leadership (related to mass tourism, with low value added and high negative externalities) or differentiation (related to the creation of unique experiences, with high value added and low negative externalities, which are understood as corresponding to the concept of sustainability).

This idea would be questioned by Butler (1999) who claims that sustainability and the principles of sustainable development should also be applied to mass tourism development processes, while concerns related to the excessive use of resources are also relevant for small scale forms of tourism in sensitive natural areas. Of particular importance for our work was the distinction proposed by this author between the impacts of tourism on sustainable development and the sustainable use of territorial resources for tourism activities (which is the perspective adopted in our analysis). This conceptualization was supported by Jafari (2001) when defining the "Knowledge Based Platform" for tourism studies who pointed out that principles of sustainability should be addressed at policy and managerial levels in all types of tourism destinations. Both authors emphasised the human dimension of sustainability and the importance of the socio-economic benefits of tourism for the host communities.

During that period, other authors stressed the importance of the uniqueness of local resources for destination differentiation along with the perception and satisfaction achieved by tourists (e.g. Kozak 1999, Buhalis 2000) while bringing attention to the sensitiveness of natural resources and the necessary limits to be imposed to their usage (Buhalis 1999, Hassan 2000). This "environmental paradox", as later defined by Williams, Ponsford (2009), implies that the production of tourism experiences depends on the exploitation of local resources, which – at the same time – must be preserved.

Synthetizing these contributions, a definition of tourism competitiveness commonly accepted in the literature was provided by Ritchie, Crouch (2003) stating that "what makes a tourism destination truly competitive is its ability to increase tourism expenditure, to increasingly attract visitors, while providing them with satisfying, memorable experiences, and to do so in a profitable way, while enhancing the well-being of destination residents and preserving the natural capital of the destination for future generations". This definition stresses the importance of growth, economic impacts, consumer satisfaction, benefits for the host community and the preservation of resources over time. Our analysis is particularly focused on the regional economic impacts and the sustainable use of resources.

Following this definition, the authors developed a comprehensive model of destination competiveness while other systematic approaches were proposed in subsequent years (e.g. Vanhove 2005, Mazanek et al. 2007). In particular, Celant (2007) focused his analysis at the regional level, while Weaver (2006), Wall, Mathieson (2006) or Sharpley (2009) offered a more clear focus on the problems of sustainability. Tsai et al. (2009) as well as Park, Jang (2014) offer detailed overviews of these formulations. Systematic approaches combining tourism competitiveness with the principles of sustainable development were proposed as policy guidelines by international institutions like UNESCO (2000, 2005), the European Commission (2007), UNWTO (2007) or the World Economic Forum (2008).

2.2 The region as the territorial level of analysis

Although some of the studies on competitiveness previously mentioned are related to empirical applications focused on particular destinations, most of them aim at international comparisons, which has lead to the creation of country rankings based on composite indicators. One example with large international recognition is the Travel and Tourism Competitiveness Index developed by the World Economic Forum (2008), which uses a very large set of quantitative indicators. Based on this index, several authors applied different methodologies in order to refine the analysis of particular aspects of tourism competiveness: Mazanek et al. (2007) selected a particular group of indicators in order to provide an explanatory model for tourism competitiveness; Navickas, Malakauskaite (2009) oriented the use of these indicators to questions related with innovation; Webster, Ivanov (2014) analysed the relation between this index and the contribution of tourism

for economic growth; Martín et al. (2017) proposed a definition of a benchmark position and country profiles regarding tourism competitiveness.

Despite the abundant number of studies on tourism competiveness at the country level, comparative analyses between regions of different countries are relatively scarce, probably due to the difficulties in obtaining relevant and comparable data. Camisón, Forés (2015) focused on the firm level and analysed how regional competitiveness influences the performance of tourism companies. Cracolici, Nijkamp (2008) related the attractiveness of Southern Italian (NUTS 2) regions with tourist satisfaction as a proxy for regional competitiveness. Closer to the purposes of our study, and focusing on a larger number of Italian NUTS 2 regions, Cuccia et al. (2016) observed that cultural and environmental regional endowment positively affect the performance of destinations, but the existence of UNESCO sites does not imply similar benefits. With a different territorial focus (centred on the destination) Cucculelli, Goffi (2016) observed that factors related to the sustainable use of resources exert positive impacts on the performance of Italian certified destinations.

Considering NUTS 2 regions as the territorial level of analysis, our study aims at offering a comprehensive overview of the relations between tourism performance and natural and cultural endowment in order to achieve policy implications both in terms of tourism development and resource management, which are addressed at a relevant institutional level for strategic orientations.

2.3 Spatial econometrics in tourism

Due to the large increment of geo-referenced statistical information recently available and the development of easy-to-use software tools, spatial econometric methodologies are only currently becoming of widespread use even though they started to be developed in the mid 20th century (Florax, Vlist 2003, Anselin 2010). In particular, our work is based on a panel data approach. This allows for the development of complex analyses of economic processes and their spatial effects while taking into consideration more information, increasing the variation and reducing the collinearity between variables resulting in more efficient estimations (Elhorst 2003, 2014).

A limited number of panel data models have been used in tourism studies, mostly over the last 10 years, as summarized by Song et al. (2012). Moreover, despite the close relationship between tourism and territory, only a few works applied spatial panel data models in tourism: Marrocu, Paci (2013) analysed the determinants of tourism flows between 107 Italian locations; Yang, Fik (2014) examined spatial spillovers and spatial heterogeneity in order to explain the variability in tourism growth across 342 cities in China; Kang et al. (2014) analysed the territorial impacts of national tourism policies in South Korea; Ma et al. (2015) focused on the spatial correlation between tourism and urban economic growth in 272 Chinese regions; Majewska (2015) applied techniques of exploratory spatial data analysis to study the inter-regional agglomeration effects in tourism activities in Poland.

Closer to the purpose of our work, Patuelli et al. (2013) examined the importance of world heritage sites on internal tourism flows in Italy, finding positive impacts on the regions where the sites are located followed by negative impacts on tourism flows in the surrounding regions, suggesting the existence of a strong competitive effect. In particular, the results obtained by Romão (2015) in a spatial analysis of the impact of natural and cultural assets on tourism demand in European regions (identifying a positive correlation between the regional endowment on natural and cultural resources and the volume of tourism demand) can be compared with those presented in this paper.

3 Data, methodologies and results

This section includes a presentation of the data and variables under analysis, a preliminary panel regression estimating the impacts of the factors considered on regional tourism competitiveness, an exploratory spatial analysis aiming at the identification of possible spatial effects and, finally, an econometric regression including spatial effects and offering an overall explanatory framework.

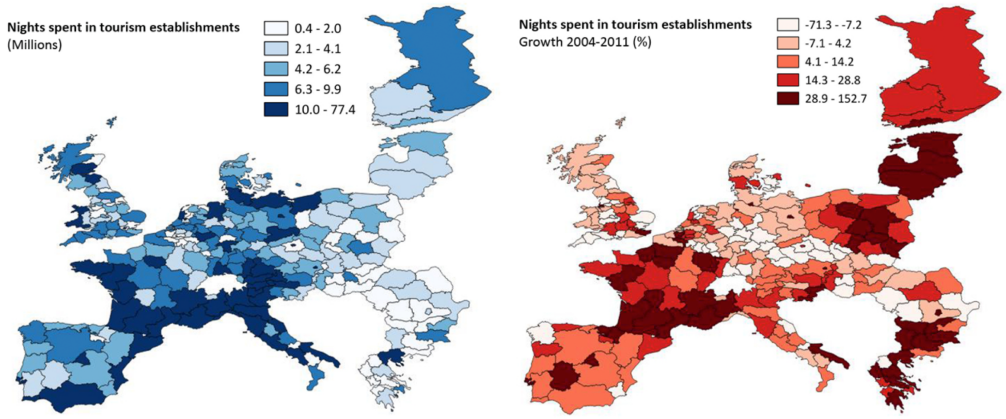


Figure 1: Regional tourism demand

3.1 Data and Variables: Tourism Dynamics, Culture and Nature

The territorial level considered in the present study is the NUTS 2 level according to the definitions of the Eurostat for the European regions. This regional scale is appropriate for the purposes of this study as it allows us to obtain relevant and comparable data and to discuss and address policy recommendations at an adequate institutional level. 237 regions (excluding islands) from Austria, Belgium, Bulgaria, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Italy, Latvia, Lithuania, Luxemburg, the Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain and the United Kingdom have been considered. Other European regions have not been included due to the lack of statistical information. The source for the geographical representation of the data in the following Figures is GISCO – Eurostat (European Commission) for the administrative boundaries of European regions, while the information related to tourism dynamics, cultural assets and natural resources has been added to that source. The maps used in this section (Figures 1 and 2) were produced with QGIS 2.4, while those used in the following section (Figure 3) were produced with GeoDa 1.6.7.

Tourism competitiveness (the dependent variable considered in the spatial panel data model to be developed below) is measured based on available Eurostat data and taking into consideration the gross value added (at current prices) by tourism activities (including wholesale and retail trade; transport; accommodation and food service activities; information and communication services) in each region. This corresponds to a broad definition of tourism services and assumes the increasing importance of information and communication technologies (ICT) for tourism. Some missing data detected for a small number of regions were computed according to the existing information, considering the registered trends immediately before and/or after the missing information.

The explanatory variables considered in the model to be presented include tourism demand (measured by the nights spent at tourism accommodation establishments, according to Eurostat, as a proxy for regional tourism attractiveness) and investment in tourism (gross fixed capital formation in the sector, also based on Eurostat data, as a proxy for the tourism related infrastructures). Natural resources are measured taking into consideration the territory of each region classified by the European Union within the Natura 2000 network (according to a harmonized set of rules and criteria, with the information available at the European Commission – DG Environment), while cultural assets are measured considering the number of World Heritage sites classified by UNESCO at the regional level (also following an international classification based on universal criteria). The evolution between 2004 and 2011 is considered for all these variables.

The information for the most recent year (2011) regarding the geographical distribution of tourism demand in European regions measured by the nights spent in accommodation establishments (hotels, holiday and other short-stay accommodation, camping grounds, recreational vehicle parks and trailer parks) is represented in the map on the left hand

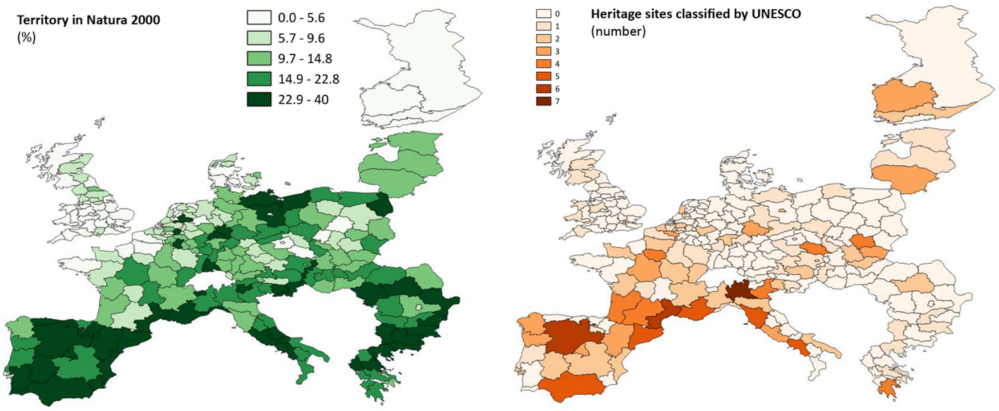


Figure 2: Natural and cultural resources

side of Figure 1, while the growth rates for this indicator between 2004 and 2011 are shown in the right hand side (the classes were created based on quintiles). The spatial pattern identified in Figure 1 reveals the importance of the Southern European regions, but the map on the right hand side of the same Figure reveals a clear shift in tourism demand with a large amount of regions located in the Eastern side of Europe (mostly Baltic, Bulgarian and Greek regions) among those with the highest growth rates registered between 2004 and 2011. However, it should be noticed that these high growth rates are also related to the low scores observed for these regions in the initial period.

The natural resources of each region have been measured taking into consideration the percentage of its territory included in Natura 2000, as a proxy for regional biodiversity. Although these protected areas are not necessarily tourism attractions, they reveal the potential attractiveness of natural resources for tourism demand in each region. In this analysis natural resources are not seen as tourism products, neither is it required that they are perceived as such (or even as protected areas) by tourists: they allow us to assess the sensitiveness and importance of regional ecosystems. The map presented on the left hand side of Figure 2 represents this information for 2011, showing the importance of Southern European regions for the biodiversity in Europe.

Finally, the number of sites classified as World Heritage by UNESCO in each region was assumed as a proxy for its cultural heritage (in a few cases, the same site is distributed along different regions and one site per region has been considered). Despite the existence of other important cultural elements (from tangible, like non-classified monuments or museums, to intangible, like local lifestyles, and including cultural events and festivals) extremely relevant for tourism attractiveness, it is not possible to have comparable quantitative information at the regional level for an international analysis. The map on the right hand side of Figure 2 represents cultural heritage as measured by the number by UNESCO sites again revealing the importance of the Southern European regions. In the same sense as it was seen for the natural resources, our analysis does not imply the utilization of these assets as tourism products or the perception of their historical importance by tourists: they represent a proxy for the richness of cultural heritage in each region.

3.2 Panel data model without spatial effects

As the purpose of this work is to analyse the effects of tourism demand, investment, natural resources and cultural assets on the regional tourism competitiveness in a large number of European regions over 8 years, a panel data model is an adequate tool. For the purposes of estimation of the models, logarithms are applied to the dependent variable (GVA in the tourism sector – "logGVA") and to some of the independent variables (such as the number of nights spent in tourism accommodation establishments – "logNIT" – and gross fixed capital formation in the tourism sector – "logINV"), in order to reduce

Table 1: Panel data model parameter estimates (individual effects; no spatial effects)

Parameters	Estimate	Std. Error	t-value
logNIT	0.146***	0.019	6.761
logINV	0.049***	0.010	4.823
NAT	-0.003***	0.001	-2.343
HERIT	0.031***	0.013	2.396

Notes: ***, **, * indicates statistical significance at a 1%, 5% and 10% significance level

the dispersion of the data. Absolute values are considered for the number of heritage sites ("HERIT") and percentages for the portion of the territory included in Natura 2000 ("NAT"). No spatial effects are considered in a first stage and the model is specified as:

$$\text{logGVA}_{it} = \beta_0 + \beta_1 \text{logNIT}_{it} + \beta_2 \text{logINV}_{it} + \beta_3 \text{HERIT}_{it} + \beta_4 \text{NAT}_{it} + u_{it} \qquad (1)$$

where the index i refers to the ith region, t is an index for the time period and u is an independent and identically distributed error term.

Although the number of periods under analysis is relatively small (8 years), the cross-sectionally augmented Im, Pesaran and Shin (IPS) test for panel unit roots (Pesaran 2007) has been applied (using the plm package in R; see Croissant et al. 2016) confirming the stationarity of the data under different test specifications allowing for individual intercepts or trends among the data, defining the number of lags of the test regression according to the Akaike Information Criteria and specifying the maximum number of lags as 2 or 4. The test statistics obtained were, respectively, -51.632 and -61.873, corresponding both to a p-value below 2.2e-16), confirming the stationarity of the data. A variance inflation test (VIF) was also computed (using the package car in R) and all the scores obtained (logNIT = 1.715; logINV = 1.807, NAT = 1.060; HERIT = 1,333) were clearly below the threshold of 5 suggested by O'Brien (2007), revealing the absence of problems of multicollinearity.

In order to choose between a fixed or a random effects model, a Hausman test has been computed with the plm package in R and its result (p-value $<$ 2.2e-16) suggested that a fixed effects model should be preferred (methodologies and techniques are described in Croissant, Millo 2008). Nevertheless, as discussed by Clark, Liner (2015), the Hausman test has important limitations for a final decision regarding the choice of a specific model, which should be grounded on theoretical assumptions about the observations. In our case – and considering the close link between the specific characteristics of the territories and tourism dynamics, it seems plausible to assume that individual regional features have specific impacts on tourism activities, also justifying the option for a fixed effects model. Thus, a panel data model with fixed effects has been estimated.

The estimation results are presented in Table 1, revealing a positive (and statistically relevant) relation between tourism GVA and all the independent variables, except for natural resources. It is possible to confirm the expected positive correlation between the existence of classified heritage sites and tourism competitiveness. It is noticeable that the variables related to tourism dynamics (demand and investment) exert a higher impact on tourism competitiveness than the impact generated by the existence of cultural assets.

The result related to the negative impact of natural resources on tourism competitiveness requires a more careful interpretation: it could be argued that this negative correlation is the expectable consequence of the fact that natural resources are measured taking into consideration the proportion of protected areas in each region (since the related conservation measures that are implied generally impose restrictions on tourism activities). Nevertheless, as observed in the previous section, these regions are mostly located in Southern Europe, which are generally places with high levels of tourism demand. In fact, a positive relation between tourism demand and the portion of the territory included in Natura 2000 had been found in a previous study on the same regions and for a similar period (Romão 2015).

Table 2: Moran I tests for spatial autocorrelation

	logGVA		logNIG		logINV		HERIT		NAT	
	2004	2011	2004	2011	2004	2011	2004	2011	2004	2011
Test Results	8.869	9.105	10.34	11.74	6.586	9.498	6.728	8.695	16.76	18.69

Thus richer natural resources are correlated to higher levels of demand but to relatively low value added. In other words, despite the apparent potential of these regions to create a differentiated tourism supply, based on their rich biodiversity within the European context, they seem to develop tourism products and services oriented to mass consumption normally with relevant negative consequences in terms of the protection of sensitive environmental resources. Therefore, a strategy of cost leadership (implying a massive use of resources with low economic impact) seems to prevail over a differentiation strategy based on the uniqueness of the places (oriented to the provision of unique experiences with protection of sensitive resources and oriented to high value added products and services).

Despite the high statistical relevance of all parameter estimates of this model, the R-squared (0,048) and the Adjusted R-squared (0,042) obtained from this regression are relatively low. Although the R-squared statistic may have important limitations when applied to time series, the results suggest that the estimation can be significantly improved. Also, the computation (with the plm package for R) of a test (Pesaran CD) for cross-sectional dependence (Pesaran et al. 2008) has lead to results (test statistic of 54.592) suggesting evidence in favour of the existence of such characteristics in the panel under analysis, opening the possibility for the existence of spatial effects.

Finally, as suggested by Clark, Liner (2015), different specifications of the model have been computed in order to confirm the stability of the results. As can be seen from Table A.1 (Appendix) the signs of the estimates for all parameters are the same independently of the type of model (fixed effects, random effects or pooling), although some differences can be observed regarding the statistical relevance of the estimates. On the other hand, a second set of models has been computed (Table A.2 in Appendix) replacing the variables that had been logarithmized (GVA, gross fixed capital formation in the tourism sector, and nights spent in accommodation establishments), which were, instead, divided by the number of residents in each region (values per habitant were obtained) in order to consider the possible effects of the dimension of the region. As can be observed the results show exactly the same signs for all estimates, but with a much lower statistical significance.

3.3 Exploratory Spatial Data Analysis: Territorial Resources and Tourism Competitiveness

In order to identify the possible existence of spatial effects, several preliminary tests were computed by using indicators of spatial autocorrelation (Anselin 2005). This methodology requires the creation of a spatial weights matrix, defining the spatial impacts of each region on its neighbours (Anselin 2005). In this case, a neighbour is defined according to the rook contiguity criteria (two regions are considered neighbours if they share a common border) and it is also assumed that spatial impacts occur, not only for immediate neighbours, but also for the "neighbours of neighbours" (second level contiguity). Additionally, it is also assumed that the impact on immediate neighbours is double than the impact on second order neighbours and that all regions have the same potential to generate spillover effects (implying that the spatial weights matrix is row normalized). The results obtained suggest that this impact matrix offers useful insights for the estimation of spatial effects.

Moran I tests for spatial autocorrelation (Anselin 2005) provide a measure for global spatial correlation between neighbours. Table 2 shows the test results obtained (using Geoda 1.6.0) for all variables included in the model and considers the first and last year of the observations. The existence of spatial correlation is suggested by the test results obtained (a pseudo significance level is computed through a random permutation procedure, recalculated 99 times in order to generate a reference distribution).

Local indicators of spatial autocorrelation have also been computed (with Geoda

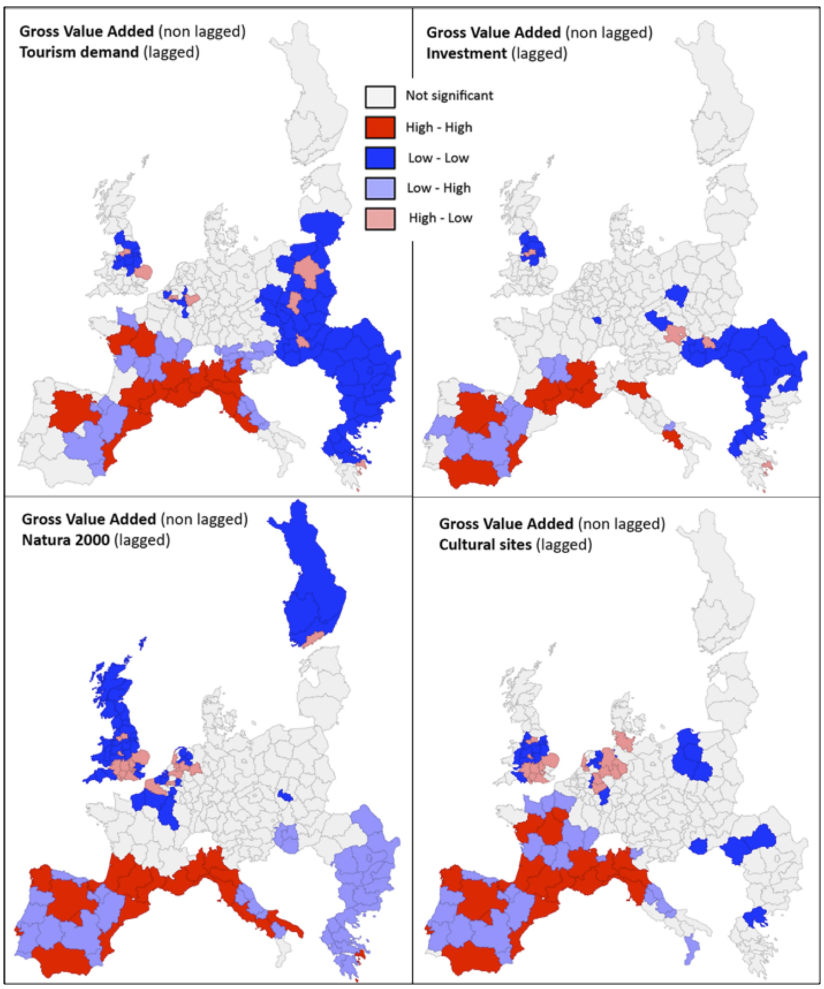


Figure 3: Local Indicators of Spatial Autocorrelation – Bivariate Analysis

1.6.0 and following the methodologies proposed by Anselin 2005) in order to generate a bivariate analysis based on the local Moran I indicators, relating a non-lagged variable (the dependent variable – GVA in tourism) with spatially lagged variables (each of the four independent variables – tourism demand, investment in tourism, natural resources and cultural resources). The maps in Figure 3 represent these spatial relations, considering a 95% significance level. Dark colours represent clusters of positively correlated regions (dark red for high values in both variables and dark blue for low values in both variables) and light colours represent negative correlation (light red for high values of the non-lagged variable and low values for the lagged variable, and light blue for the inverse situation).

The first map (top-left) reveals a cluster of high values for tourism GVA and high tourism demand in southern western regions (dark red), while clusters of low values for both variables are located in the east side of Europe (dark blue). Nevertheless, the existence of a large number of southern regions (mostly in Spain, France and Italy) where low value added in tourism is spatially correlated with high tourism demand (light blue) is also noticeable, suggesting that tourism is possibly based on low value added products and services. A very similar pattern is observed for the second map (top-right), revealing that a large number of regions from Southern Europe (mostly concentrated in Spain) register high levels of investment in tourism while achieving relatively low levels of GVA.

The combination of these results (high tourism demand and high investments in the tourism sector) suggests a high mobilization of regional resources for tourism. However they generate a relatively low value added, which can be related to a low productivity in the utilization of these resources (nevertheless, it should be noted that this study does

not address the specific question of productivity and does not provide a measure for the relation between the output of tourism activities and the necessary inputs for their provision). This tendency is even more marked when we observe the relation between GVA and natural resources (down-left) or cultural heritage (down-right). In the second case (cultural heritage), we can observe that for a large number of regions from Portugal, Spain, France and Italy low gross value added is correlated with a high number of cultural sites classified by UNESCO. On the other hand, when natural resources are taken into consideration, this tendency is also observable for Greek and Bulgarian regions (although it does not happen in France, suggesting that higher value added is achieved with nature oriented tourism in French regions).

Generally, these observations confirm the results obtained from the panel data model previously estimated: the negative correlation between natural resources and the GVA generated by tourism activities is not related to a low level of tourism demand in protected areas, but to the supply of massive, low value added tourism products and services in these regions. Even if some of them (located in coastal areas) register high levels of tourism GVA, a large number of territories (mostly those without direct connection to the sea) do not achieve good performance in terms of GVA despite the high tourism demand.

3.4 Spatial Econometric Analysis

The final step of this work is the computation of a spatial regression model by adding the spatial effects explicitly to the panel data model presented in (1). The existence of a spatially lagged endogenous variable (included as one more explanatory variable and capturing potential endogenous interaction effects) and spatial effects in the error term (a spatial multiplier that captures un-modelled spatial effects expressed in the interaction among the error terms) will be tested before estimation. In a general form, a space-time panel data model with spatial effects among the dependent variables and the error term can be specified as:

$$\begin{aligned} Y_{it} &= \rho(WY)_{it} + X'_{it}\beta + u_{it}, \\ u_{it} &= \lambda(Wu)_{it} + \epsilon_{it}, \qquad\qquad i = 1, \ldots, N \ \text{ and } \ t = 1, \ldots, T \end{aligned} \tag{2}$$

where

- Y_{it} represents the log of tourism GVA in region i at time t.
- X_{it} corresponds to a 4x1 vector of independent variables for region i at time t, namely:
  - the number of nights in tourism establishments;
  - the gross fixed capital formation in tourism;
  - the percentage of the territory classified in Natura 2000;
  - and the number of sites classified as World Heritage by UNESCO.
- W is a nonnegative $N \times N$ matrix of known constants describing the spatial impacts; the element w_{ij} indicates the intensity of the relationship between cross sectional units i and j and the diagonal elements are set to zero because no region can be its own neighbour;
- WY represents the endogenous interaction effects among the dependent variables;
- Wu shows the interaction effects among the disturbance terms of the different units;
- ρ is the spatial autoregressive coefficient;
- λ the spatial autocorrelation coefficient;
- i is an index for the regions and t is an index for the time period.

Table 3: Parameter estimates of the spatial lag and spatial error model with fixed effects

Parameter	Estimate	Std.Error	t-value
Intercept	-0.330***	0.028	3.843
logNIT	0.252***	0.014	17.787
logINV	0.655***	0.013	50.138
NAT	-0.019***	0.001	-14.940
HERIT	-0.005	0.008	-0.675
Spatial autoregressive coefficient (ρ)	0.402***	0.040	10.067
Spatial autocorrelation coefficient (λ)	0.109***	0.028	3.843

Notes: Note: ***, **, * indicates statistical significance at a 1%, 5% and 10% significance level.

The tests for the existence of spatial effects (Baltagi et al. 2003, 2007) were performed using the splm package in R (Milo, Piras 2012), aiming to identify whether the potential spatial effects are related to regional effects within the dependent variable (a spatial lag model, with $\lambda = 0$) and/or more general effects identified in the spatial distribution of the error terms (spatial error model, with $\rho = 0$). The score of 1469.973 obtained for the LM (Lagrange Multiplier) test implies the rejection of the null hypothesis of no random effects and no spatial autocorrelation (H_0: $\lambda = \rho = 0$) and suggests the existence of spatial effects related to the dependent variable and/or the spatial correlation among the error terms (alternative hypothesis is that at least one component is not zero). In fact, the Moran I test computed in the previous section had already shown the existence of spatial effects (implying $\rho \neq 0$).

The Baltagi, Song and Koh's SLM1 marginal test evaluates the inexistence of autoregressive spatial effects (H_0: $\rho = 0$), assuming that no spatial effects exist in the error term ($\lambda = 0$); the score of 0.0187 with a p-value of 0.9851 implies non-rejection of the null hypothesis. Conversely, Baltagi, Song and Koh's SLM2 marginal test, tests the null hypothesis of no spatial effects in the error term (H_0: $\lambda = 0$) assuming no autoregressive spatial effects ($\rho = 0$); the score of 0.00847 with a p-value of 0.9933 also implies non-rejection of the null hypothesis. Thus, the inexistence of one type of spatial effects also implies the inexistence of the other. Finally, applying Baltagi, Song and Koh's conditional test, LMλ, for no regional effects expressed in the error term (H_0: $\lambda = 0$), independently of the value of ρ, a score of 37.1293 was obtained leading to the rejection of the null hypothesis. Thus, it is possible to conclude for the existence of spatial autocorrelation effects in the error term.

The existence of both types of spatial effects leads to the computation of a general spatial Cliff-Ord type model (Cliff, Ord 1981), including a spatially lagged dependent variable and a spatially autocorrelated error term. Finally, the results obtained for a spatial Hausmann test lead us to opt for a fixed effects model.

This spatial lag and spatial error model [with: Y =logGVA and X =(logNIT, log INV, NAT, HERIT)] is defined according to expression (2) previously presented, considering fixed individual effects and requiring a specification of the disturbances assuming that spatial autocorrelation applies to both the individual effects and the error term (with the transformation proposed by Kapoor et al. 2007, for the disturbance term following a first order spatial autoregressive process – "kkp" type). Table 3 presents the results obtained based on maximum likelihood estimation:

Although the computation of the R-square statistic is not possible in a spatial context, other results (also with important limitations, as discussed by Elhorst 2014) reveal a relevant increase in the adjustment regarding the model without spatial effects. The computation of a squared correlation coefficient between actual and fitted values (proposed by Elhorst 2014) has led to a result of 0.833 (0.320 for the model without spatial effects previously presented), while the computation of a Pseudo R-squared based on the quotient between the variance of the estimations and the variance of the actual values (proposed by Anselin, Lozano-Gracia 2008) has led to a result of 0.883 (0.830 for the model without

spatial effects). Even if the measures of goodness of fit have important limitations in both cases, the model with spatial effects clearly performs better than the model without spatial effects.

Comparing the estimated parameters with those obtained from the model without spatial effects, the same type of relations (expressed in the sign of the correlations) between the dependent variable and the independent variables were identified although the impact of cultural assets loses statistical significance when spatial effects are included. This model confirms the expected positive correlation between tourism GVA, tourism demand and investment in the tourism sector, but in this case the impact of investment is higher than the impact of tourism demand. Possibly, a part of the regional dynamics (linked to tourism demand) is now captured by spatial effects associated to the lagged variable (tourism GVA), while the impact of cultural assets can eventually be captured by unmodelled spatial effects related to the error term. The model also confirms the negative correlation between regional natural resources and tourism GVA previously observed, which is the most important result arising from this analysis.

The existence of spatial effects among regions is also clear. The spatial effects identified in the space-time model reveal the existence of spillover effects (expressed in the positive value of the spatial autoregressive coefficient, showing that tourism dynamics in one region has positive consequences on the contiguous regions) and also unmodelled effects (expressed in the positive value of the spatial autocorrelation coefficient). Although it is clear that a major part of these effects are captured by the spatial distribution of the dependent variable (with a much larger estimated parameter), the existence of spatial effects in the distribution of the error terms suggests that other type of variables can be included in further works in order to increase the explanatory power of the model.

4 Discussion and Concluding Remarks

A first important conclusion of this study is the confirmation of the useful contribution of spatial analysis in tourism studies with a clear impact on the goodness of fit of the econometric model and the identification of spatial patterns in tourism activities and its determinants. It was possible to conclude from our exploratory spatial analysis that the impacts of the determinants of competitiveness taken into consideration differ across the territorial units despite the existence of a general trend identified by the econometric model. This is the first contribution of this work, leading from a policy point of view to the idea that the implementation of guidelines to improve tourism competitiveness must take into account the specific territorial conditions.

The results of this spatial analysis also imply that contemporary regional tourism dynamics is related, not only to regional resources and conditions, but also to the dynamics observed in neighbouring regions. This also has clear implications for tourism policies suggesting that local resource management, promotional strategies, transport systems or accommodation provision can be more efficiently planned if there is some collaboration among clusters of regions with similar characteristics. In fact, this type of complementarity between regions is possible to observe in many parts of Europe, even belonging to different countries as can be observed in e.g. mountain areas like the Alps (Switzerland and Italy) or the Pyrenees (Spain and France) or along major rivers (like the Danube).

The analysis of the determinants of regional competitiveness developed through the computation of a spatial panel data model confirmed the expected positive correlations between GVA generated by tourism activities, regional tourism demand and investment in the tourism sector. Nevertheless, the results also showed a negative relation between tourism GVA and the existence of natural resources. Although this negative correlation could be linked to the type of data used in the model (suggesting that it could be related to protective measures implemented in these areas), it was also possible to observe that regions with more protected areas are located in Southern European areas with high levels of tourism demand where mass tourism prevails.

A second contribution of this work for the existing literature is the identification of this negative correlation between the existence of rich natural resources and tourism competitiveness in European regions, which was only possible through the combined

analysis of the econometric model (providing an overall general explanation) and the exploratory spatial analysis (identifying different spatial patterns and highlighting the role of Southern European territories in this context). In fact, the indicators of spatial autocorrelation used for the exploratory spatial analysis revealed the existence of a large number of regions from Southern Europe where abundant natural resources coexist with high levels of tourism demand and low value added in the tourism sector. Thus, the results suggest that massive tourism related to natural resources tends to generate less positive impacts on the regional GVA, despite its potential negative impacts on ecosystems and landscapes.

This analysis reveals an unsustainable process of tourism development for these regions apparently following a cost leadership competitive strategy based on low prices. Instead, taking into consideration their richness in terms of natural and cultural assets, a strategy of differentiation aiming at the provision of unique experiences based on the specific territorial resources could lead to a more sustainable form of tourism development. This would reinforce the linkage with other local economic activities with larger impacts on regional development and higher protection of sensitive resources. In fact, good practices related to this kind of utilization of natural resources for the creation of high value tourism products and services can already be found in many natural parks all over the world including Europe, while countries like Australia or New Zealand tend to give very high importance to these services within their tourism activities.

Finally, it is important to notice that the increasing amount of geo-referenced information related to tourism opens new opportunities for the application of spatial analysis techniques in order to identify spatial patterns of tourism development. Questions related to the effective usage of natural or cultural resources as tourism products or a more detailed analysis of tourism infrastructures can be integrated into similar models in the future, along with the consideration of other determinants of tourism competitiveness (marketing, management, planning, etc.). Another possible development of this work relates to the scale of analysis, given that NUTS 2 regions can include different tourism destinations within the same territory. The NUTS 3 level can be more appropriate for this purpose when comparable relevant statistical information is available.

References

Anselin L (2005) Exploring spatial data with GeoDa™: A workbook. Center for Spatially Integrated Social Science, Santa Barbara http://la1.rcc.uchicago.edu/media/-geoda_files/docs/geodaworkbook.pdf

Anselin L (2010) Thirty years of spatial econometrics. *Papers in Regional Science* 89: 3–25. CrossRef.

Anselin L, Lozano-Gracia N (2008) Errors in variables and spatial effects in hedonic house price models of ambient air quality. *Empirical Economics* 34: 5–34. CrossRef.

Baltagi S, Jung B, Koh W (2003) Testing panel data regression models with spatial error correlation. *Journal of Econometrics* 117: 123–150. CrossRef.

Baltagi S, Jung B, Koh W (2007) Testing for serial correlation, spatial autocorrelation and random effects using panel data. *Journal of Econometrics* 140[1]: 5–51. CrossRef.

Buhalis D (1999) Limits of tourism development in peripheral destinations: problems and challenges. *Tourism Management* 20: 183–185

Buhalis D (2000) Marketing the competitive destination of the future. *Tourism Management* 21: 97–116. CrossRef.

Butler R (1999) Sustainable tourism: A state-of-the-art review. *Tourism Geographies* 1[1]: 7–25. CrossRef.

Camisón C, Forés B (2015) Is tourism firm competitiveness driven by different internal or external specific factors? New empirical evidence from Spain. *Tourism Management* 48: 477–499

Celant A (2007) *Global Tourism and Regional Competitiveness.* Patron, Bologna

Clark T, Liner D (2015) Should I use fixed or random effects? *Political Science Research and Methods* 3[2]: 399–408. CrossRef.

Cliff A, Ord J (1981) *Spatial Processes, Models and Applications.* Pion, London

Cracolici M, Nijkamp P (2008) The attractiveness and competitiveness of tourist destinations: A study of southern Italian regions. *Tourism Management* 30: 336–344. CrossRef.

Croissant Y, Millo G (2008) Panel data econometrics in R: The plm package. *Journal of Statistical Software* 27[2]: 1–43. CrossRef.

Croissant Y, Millo G, Tappe K, Toomet O, Kleiber C, Zeileis A, Henningsen A, Andronic L, Schoenfelder N (2016) Package "plm". CRAN, https://cran.r-project.org/web/-packages/plm/plm.pdf

Cuccia T, Guccio G, Rizzo I (2016) The effects of UNESCO world heritage list inscription on tourism destinations performance in Italian regions. *Economic Modelling* 53: 494–508. CrossRef.

Cucculelli M, Goffi G (2016) Does sustainability enhance tourism destination competitiveness? Evidence from Italian destinations of excellence. *Journal of Cleaner Production* 111: 370–382. CrossRef.

Elhorst J (2003) Specification and estimation of spatial panel data models. *International Regional Science Review* 26: 223–243. CrossRef.

Elhorst J (2014) *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels.* Springer, New York

European Commission (2007) Agenda for a sustainable and competitive European tourism. European Commission, Luxembourg. http://eur-lex.europa.eu/legal-content/-EN/TXT/?uri=celex:52007DC0621

Florax R, Vlist A (2003) Spatial econometric data analysis: Moving beyond traditional models. *International Regional Science Review* 26[3]: 23–243. CrossRef.

Hassan S (2000) Determinants of market competitiveness in an environmentally sustainable tourism industry. *Journal of Travel Research* 38[3]: 239–245. CrossRef.

Jafari J (2001) The scientification of tourism. In: Smith V, Brent M (eds), *Hosts and Guests Revisited: Tourism Issues of the 21st Century*. Cognizant Communication Corporation, New York

Kang S, Kim J, Nicholls S (2014) National tourism policy and spatial patterns of domestic tourism in South Korea. *Journal of Travel Research* 53[6]: 791–804. CrossRef.

Kapoor M, Kelejian HH, Prucha IR (2007) Panel data model with spatially correlated error components. *Journal of Econometrics* 140[1]: 97–130. CrossRef.

Kozak M (1999) Destination competitiveness measurement: Analysis of effective factors and indicators. European Regional Science Association Conference Papers, Dublin, http://www-sre.wu-wien.ac.at/ersa/ersaconfs/ersa99/Papers/a289.pdf

Ma T, Hong T, Zhang H (2015) Tourism spatial spillover effects and urban economic growth. *Journal of Business Research* 68[1]: 74–80. CrossRef.

Majewska J (2015) Inter-regional agglomeration effects in tourism in Poland. *Tourism Geographies* 17[3]: 408–436. CrossRef.

Marrocu E, Paci R (2013) Different tourists to different destinations – evidence from spatial interaction models. *Tourism Management* 39: 71–83. CrossRef.

Martín J, Mendoza C, Román C (2017) A DEA travel-tourism competitiveness index. *Social Indicators Research* 130[3]: 1–21. CrossRef.

Mazanek J, Wober K, Zins A (2007) Tourism destination competitiveness: From definition to explanation? *Journal of Travel Research* 46: 86–95. CrossRef.

Milo G, Piras G (2012) splm: Spatial panel data models in R. *Journal of Statistical Software* 47[1]: 1–38. CrossRef.

Navickas V, Malakauskaite A (2009) The possibilities for the identification and evaluation of tourism sector competitiveness factors. *Engineering Economic* 1[61]: 37–44

O'Brien R (2007) A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity* 41[5]: 673–690. CrossRef.

Park J, Jang S (2014) An extended gravity model: Applying destination competitiveness. *Journal of Travel & Tourism Marketing* 31[7]: 799–816

Patuelli R, Mussoni M, Candela G (2013) The effects of world heritage sites on domestic tourism: A spatial interaction model for Italy. *Journal of Geographical Systems* 15: 369–402. CrossRef.

Pesaran MH (2007) A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics* 22[2]: 265–312. CrossRef.

Pesaran MH, Ullah A, Yamagata T (2008) A bias-adjusted LM test of error cross-section independence. *The Econometrics Journal* 11[1]: 105–127. CrossRef.

Poon A (1994) The 'new tourism' revolution. *Tourism Management* 15[2]: 91–92. CrossRef.

Porter M (1985) *Competitive advantage – creating and sustaining superior performance*. The Free Press, New York

Porter M (2003) The economic performance of regions. *Regional Studies* 37[6/7]: 549–578. CrossRef.

Ritchie J, Crouch G (2003) *The competitive destination: A sustainable tourism perspective.* CABI International, Oxfordshire. CrossRef.

Romão J (2015) Culture or nature: A space-time analysis on the determinants of tourism demand in European regions. Discussion papers spatial and organisational dynamics 14, http://www.cieo.pt/discussionpapers/discussionpapers14.pdf, Research Centre for Spatial and Organizational Dynamics, Faro

Sharpley R (2009) *Tourism Development and the Environment: Beyond Sustainability?* London

Song H, Dwyer L, Cao Z (2012) Tourism economics research: A review and assessment. *Annals of Tourism Research* 39[3]: 1653–1682. CrossRef.

Tsai H, Song H, Wong K (2009) Tourism and hotel competitiveness. *Journal of Travel and Tourism Marketing* 26[5]: 522–546. CrossRef.

UNESCO (2000) *Sustainable Tourism and the Environment.* UNESCO, Paris

UNESCO (2005) *World Heritage Centre – Sustainable Tourism Programme.* UNESCO, Paris

UNWTO (2007) *Practical Guide to Tourism Destination Management.* UNWTO, Madrid

Vanhove N (2005) *The Economics of Tourism Destinations.* Elsevier, Oxford. CrossRef.

Wall G, Mathieson A (2006) *Tourism: Change, Impacts and Opportunities.* Pearson, Essex

Weaver D (2006) *Sustainable tourism: Theory and practice.* Elsevier, Oxford. CrossRef.

Webster C, Ivanov S (2014) Transforming competitiveness into economic benefits: Does tourism stimulate economic growth in more competitive destinations? *Tourism Management* 40: 137–140. CrossRef.

Williams P, Ponsford I (2009) Confronting tourism's environmental paradox: Transitioning for sustainable tourism. *Futures* 41: 396–404. CrossRef.

World Commission on Environment and Development (1987) *Our Common Future.* Oxford University Press, New York

World Economic Forum (2008) *The Travel and Tourism Competitiveness Report 2008.* World Economic Forum, Geneva

Yang Y, Fik T (2014) Spatial effects in regional tourism growth. *Annals of Tourism Research* 46: 144–162. CrossRef.

A Appendix

Table A.1: Panel estimations

	Fixed Effects Est.	Random effects Est.	Pooling effects Est.
Intercept		2.978***	0.686***
NAT	-0.003*	-0.007***	-0.021***
HERIT	0.031*	0.062	0.003
logINV	0.049***	0.130***	0.686***
logNIG	0.146***	0.316***	0.236***
Adj. R2	-0.090	0.236	0.802

In this case, logarithms were applied to the variables Tourism Gross Value Added (dependent variable), Gross Fixed Capital Formation in tourism (logINV) and nights spent in tourism accommodation establishments (logNIG).

Table A.2: Alternative panel estimations

	Fixed Effects Est.	Random effects Est.	Pooling effects Est.
Intercept		4180.750***	2800.108***
NAT	-1.094	-6.397	-83.236***
HERIT	25.354	69.189	62.367
INVpc	0.071	0.196***	3.677***
NIGpc	100.813***	127.257***	61.196***
Adj. R2	-0.120	0.052	0.553

In this case, the values for the variables Tourism Gross Value Added (dependent variable), Gross Fixed Capital Formation in tourism (INVpc), and nights spent in tourism accommodation establishments (NIGpc), were divided by the number of residents (per capita), in order to consider the dimension of the regions. Nevertheless, the results obtained were much less significant.

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 3, 2017, 19–38
DOI: 10.18335/region.v4i3.189

journal homepage: region.ersa.org

A Tale of Hidden Cities

Anastasia Panori[1]

[1] Panteion University of Social and Political Sciences, Athens, Greece (email: an.panori@panteion.gr)

Received: 8 February 2017/Accepted: 15 June 2017

Abstract. Hidden cities within a city? A large trending literature concerning urban and suburban poverty concentration patterns has been developed during the last decade. However, there are few cases where adequate data exist at a low spatial level, allowing the exploration of such socio-spatial phenomena. This paper seeks to investigate the structure and evolution of poverty within urban and suburban space, under a multidimensional framework, during a period of extended economic transformation. This paper uses the metropolitan area of Athens as its main case study, for which data at a municipal level exist, allowing the calculation of the Multidimensional Poverty Index (MPI) for the years 2006 and 2011. An extended cluster analysis, based on the calculated MPI values, results in the specification of three poverty clusters within Athens. The decomposition of the MPI index into its main dimensions highlights any existing differences between the structural and behavioural characteristics of each of them. The results indicate that there is a clear spatial concentration of poverty in the west suburban areas of Athens. The urban core of the city is characterised by middle-income municipalities, whilst the north-east and the south-east suburban areas experience low-poverty indicators. Finally, the results suggest that during the period under investigation more deprived areas were affected the most.

Key words: urban poverty; multidimensional poverty index; dynamic decomposition; Athens

1 Introduction

> *"Thus, the spatial order of human existence arises from the (social) production of space, the construction of human geographies that both reflect and configure being in the world."*
>
> E. W. Soja "Postmodern Geographies" (1989)

The investigation of socio-spatial dynamics in an urban context has long been a productive field of reflection and research by social geographers and urban poverty analysts. The existence of a socio-spatial dialectic within urban agglomerations, is one of the most commonly accepted perspectives, highlighting a bidirectional effect between cities and people who live in them. Until recently, poverty researchers have tried to investigate social phenomena by treating space as a "neutral" component of their analysis. However, it is essential to notice that "the importance of space lies in understanding it as a material product of social relations, a manifestation of social relations, and a social relation itself" (Gotham 2003, p. 724).

The fact that spatial distance follows social distance has been strongly highlighted in the literature (Duncan, Duncan 1955, Park et al. 1925). Massey, Denton (1993) indicate that poverty concentration and its variation through space are closely related to changes in the socio-economic composition of areas. This spatial segregation process has been also evaluated by Massey (2007), who points out that the conceptual - social in our case - categorization of space is a natural human tendency, based on attained and intrinsic characteristics of space. These spatial characteristics may include various parameters affecting poverty concentration, such as economic and social formation channels. As Musterd et al. (2015) state "structural theory suggests that there are at least four key factors shaping socioeconomic segregation: social inequalities, changing economic structures and levels of global connectedness, welfare regimes, and housing systems" (p. 2).

Starting from the economic driven forces related to poverty concentration, economic restructuring processes are crucial parameters responsible for reinforcing this phenomenon. Jargowsky (2002) highlights the importance of the labor market structure as an explanatory variable for poverty concentration in urban areas. The transition from a Fordist to a post-Fordist economic structure in cities has resulted in the creation of many unemployed workers, most of them characterized by low level of education or skills, that do not properly fit the new jobs, created under this post-industrial framework (Sassen 1991, Musterd 2005, Scott, Storper 2015). Given this fact, minorities with a low level of education have gradually been isolated, both in economic and social terms, giving rise to poverty concentration phenomena in cities (Cooke, Marchant 2006). Thus, the spatial isolation has reduced their integration opportunities through the labor market channel, fostering social segregation, especially in manufacturing cities or cities that experience economic restructuring processes (Musterd 2005). Under this scope, Andersson, Hedman (2016) argue that economic recession periods should also be received as economic transformation processes, due to the geographically uneven expansion of unemployment they cause.

The existence of a selective migration phenomenon, occurring towards or within the metropolitan areas, is also considered to be an important factor affecting the concentration of poverty. Flows of poor people towards the urban core of cities are driven by increased integration opportunities that exist there. Alongside with this phenomenon, counterweight selective migration movements of non-poor out of the central city area also exist (Jargowsky 2002). As middle or high-income residents decide to leave poor areas, they are being replaced by in-movers with lower socio-economic status (Andersson, Hedman 2016). It becomes evident that both cases lead to an increased concentration of poverty within the urban core.

However, in the case of the largest US metropolitan areas, a diversified behavior was observed during the 1990s. Several studies (Berube, Frey 2002, Jargowsky 2003, Kingsley, Pettit 2003, Cooke, Marchant 2006, Allard, Paisner 2016) have indicated a decline in urban core poverty during this period, accompanied with a rise in suburban poverty. Cooke, Marchant (2006) point out that increases in urban core poverty concentrations are related to the general health of the metropolitan economy, whereas respective changes in the concentration of suburban poverty relate to rapid population growth. Moreover, Cooke (2010) states that a potential rise in suburban poverty could eventually produce negative neighborhood effects in suburban areas, leading consequently to a re-concentration of poverty in the inner-ring[1] suburbs.

Under this framework, this paper tries to shed light on the developmental process of poverty concentration within the metropolitan area of Athens between 2006 and 2011. Is there any pattern of poverty concentration within this metropolitan area? Is there any evidence supporting the existence of growing suburban poverty? Given the fact that throughout this period extended economic changes have been taking place in Greece, it is crucial to investigate whether poverty concentration has changed.

Another novelty of this study, is the choice of the Multidimensional Poverty Index

[1] Cooke, Marchant (2006) define three types of regions in a metropolitan area: a) urban core; b) inner ring; and c) outer ring. Specifically, inner ring is defined as regions that are not identified as part of the urban core, with greater than 400 1950–1969 housing units per square miles, plus any continuous tract containing more than 200 1950–1969 housing units per square mile and a population density of at least 1,000 people per square mile.

(MPI) as the main methodological framework for approximating poverty instead of traditional income-based measures, as it is considered to be one of the most appropriate measures for comparing uneven development patterns between areas. The MPI index has been calculated at a municipal level, allowing to perform a cluster analysis within the metropolitan area of Athens. The cluster analysis reveals existing hidden cities within the city of Athens, that seem to follow a central-city/suburban model of poverty concentration.

The outline of this paper is the following: in Sections 2 and 3 a related literature review is presented regarding urban poverty studies for the US and EU countries, including the case of Athens. Section 4 provides a detailed analysis of the main methodological contexts that were adopted in the research. Finally, Sections 5 and 6 provide a detailed analysis of the results, along with some concluding remarks.

2 Related Literature

During recent years, a series of studies have sought to investigate the spatial organization of poverty within metropolitan areas. Large urban agglomerations, located in highly developed countries, have most of the times been in the spotlight. Most studies explore poverty concentration in large US metros (Jargowsky 1996, 2002, 2013, Berube, Frey 2002, McMullen, Smith 2002, Kingsley, Pettit 2003, Cooke, Marchant 2006, Cooke 2010, Bischoff, Reardon 2013, Kneebone, Berube 2013, Kneebone 2014, Cooke, Denton 2015, Allard, Paisner 2016), due to higher data availability at low spatial levels, such as municipalities and neighborhoods. Most of these studies indicate a shift in the spatial location of poverty across many US metropolitan areas. The pattern of an increasing central city poverty concentration that appeared between 1970 and 1990, was replaced by a rising suburban poverty concentration in the following two decades. The number of poor persons in the suburban areas almost doubled between 1990 and 2014, illustrating a growth rate higher than the corresponding population growth (Allard, Paisner 2016), and thus leading to poverty concentration in those areas.

Nonetheless, a significant part of the literature focuses on the investigation of poverty and deprivation within EU urban areas (Pinch 1993, Pacione 2004, Hunter 2014, Musterd 2005, Hamnett 2003b,a, Lee et al. 2014, Musterd et al. 2015, Andersson, Hedman 2016). European cities, in general, indicate lower levels of socio-economic segregation, when compared to cities in the US (Musterd 2005, Musterd et al. 2015). However, income disparities between different socio-economic groups seem to have risen in many cases throughout the last decade, leading to higher concentrations of urban poverty (Musterd et al. 2015).

In terms of measurement, it is important to notice that most of the existing studies use monetary (Green 1998, Reardon, Bischoff 2011, Watson 2009), racial or gender criteria (Bischoff, Reardon 2013, Jargowsky 2013, Kneebone 2014) for evaluating socio-economic segregation within cities. However, Musterd (2005) points out that in the European case, socio-economic inequality should not be encompassed strictly in one standard, such as income, but instead, various indicators of poverty should be used, in order to capture the multidimensional aspects of this phenomenon. To the authors' knowledge, the number of existing studies using alternative measures of poverty concentration within urban areas is extremely limited, due to lack of appropriate disaggregated data. However, the need for calculating and mapping various aspects of poverty at a small-area level starts to become essential, as understanding spatial distribution of poverty in modern metropolitan areas has turned out to be crucial for defining the main challenges of urban growth in the future (Allard, Paisner 2016).

In the case of Britain there has been a great effort to map multidimensional deprivation within urban areas, starting from the work of Noble et al. (2006) and moving on to the creation of the English Indices of Deprivation (2004, 2007, 2010). Highly detailed maps of London, for the years 2004, 2007 and 2010, illustrate the distribution of the Multiple Deprivation index, highlighting the existence of possible clusters within the city. Harris, Longley (2004) also try to locate possible deprivation clusters within the city of London, by combining demographic British census data with lifestyle indicators, supplied by a

UK-based data warehouse company. Referring to London, Orford (2004) attends to reveal spatial clusters of concentrated poverty and affluence. His results suggest that clusters of affluent areas tend to be more stable through time, whereas clusters including deprived units change more easily over time. In terms of suburban poverty, a recent study by Hunter (2014) indicates that poverty is becoming a problem for suburbs, as the results for England and Wales illustrate a converging trend between the urban cores and the suburbs, in terms of poverty concentration.

Based on the previous, it becomes evident that a comprehensive analysis trying to shed light on the evolution of poverty concentration within a metropolitan area, should include insights referring not only to various structural aspects of poverty, but also to its behavioural characteristics through time.

3 The case of Athens

In general, the metropolitan area of Athens covers the largest part of the greater Region of Attica in Greece. Until recently, it has exhibited high development rates, especially in the 1990s and the early 2000s, just before the economic crisis period. According to Pantazis, Psycharis (2016), the evolution of income within the region of Attika during the 2000s can be divided in three discrete periods referring to the last decade. These include: (a) the period 2003-2008, when a general uprising trend in income was recorded; (b) the period 2008-2010, characterized by an income stagnation and at the same time, a deterioration of social conditions; and (c) the period 2011-2013, where the effects of the 2008 economic crisis started to become evident, affecting income distribution.

Regarding spatial inequality, there seems to be an East-West division of Athens, in terms of residential inequality (Maloutas 2001, Kalogirou 2011), as well as income distribution (Pantazis, Psycharis 2016). A comprehensive presentation of the evolution of socio-economic segregation in Athens has been made by Maloutas (2015), illustrating the evolution process that took place, starting before the 1970s and moving on until the 2000s. According to it, there is a discrete spatial socio-economic distribution pattern within Athens, in which high income areas are concentrated in the north-eastern and southern-eastern parts, whilst low income areas are traditionally located in its western areas. Table 1 illustrates a brief description of the main findings presented in Maloutas (2015), regarding the spatial segregation process within Athens during recent decades.

In terms of urban core and suburban development, the pre-1970s in-flow pattern towards the city center seems to be replaced, between 1970 and 1990, by a movement of middle- and high-social classes towards suburban areas. This trend is followed by a significant arrival of immigrants during the 1990s, who are mostly settled within the urban core of Athens, where they could find affordable housing. This fact has led to a class desegregation period within the city center (Maloutas 2015). During the 2000s, there were no significant changes in the existing spatial segregation patterns, despite the high level of social mobility in working-class areas. This might be due to several reasons including family solidarity networks, importance of spatial proximity with family and the fact that parental property is often located in the same area.

Based on the abovementioned analysis, we explore poverty concentration within the metropolitan area of Athens, for the years 2006 and 2011, through identifying discrete clusters with different socio-economic characteristics. Instead of using monetary poverty indicators, we choose to adopt the multidimensional poverty framework (MPI index). Finally, the choice of the years 2006 and 2011 was based on the ability to compare the results between two different time frames: one related to a period of relative economic stability and growth, whilst the other to a period experiencing economic restructuring and recession. It is essential to investigate whether poor areas were affected to a wider extent during this transition.

4 Data and methodology

This section illustrates the methodological framework for calculating and decomposing MPI index, as well as performing cluster analysis based on the attained results. Data used

Table 1: Evolution of the spatial socio-economic segregation process within the metropolitan area of Athens (Maloutas 2015)

Period	Description of segregation process
Before 1970s	Rapid urbanization, especially in the urban core of Athens. Decline in living conditions in the urban core.
1970–1990	Geography of social segregation started to change (Maloutas 2000) Sub-urbanization trend. People belonging to high and middle-class start to move in the suburbs, mostly to north-east and south-east areas. Suburban growth period. Lower-classes traditionally remain settled in the western parts of the metropolitan area of Athens.
1990–2000	Large percentage of immigrants leads to lower levels of class segregation within the urban core, as immigrants can find affordable apartments only at the city center of Athens (Maloutas 2007, Maloutas et al. 2012)
During 2000s	No essential changes in the traditional social division of Athens, between east and west. High social mobility movements in working-class suburbs, not followed by high levels of residential mobility (Maloutas et al. 2006). Spatial entrapment of socially mobile groups due to family solidarity networks, importance of spatial proximity with family and parental property located in the same area (Maloutas 2004).

as input in this paper include parts of the dataset developed for the metropolitan area of Athens by Panori et al. (2017), which includes the calculated values of the MPI index at a municipal level. A further exploitation of these results through this study includes: a) the definition of an optimal number of poverty clusters within Athens, alongside with the presentation of their main socio-economic characteristics; b) the decomposition of the regional MPI index into its main dimensions for each of them; and c) the dynamic decomposition of all derived indices between 2006 and 2011. All these processes will help to identify any existing structural and behavioral differences of poverty within the city of Athens.

Starting from the empirical measurement of multidimensional poverty, the Alkire-Foster methodology has been used to conceptualise this theoretical framework (Alkire, Foster 2011a,b). Following the Apablaza, Yalonetzky (2013) notation, the multidimensional headcount ratio measuring the percentage of population being multidimensionally poor in period t is defined as follows:

$$H(t) \equiv \frac{1}{N^t} \sum_{n=1}^{N} I(c_n^t \geq k) \tag{1}$$

where N is the number of individuals, k is the multidimensional cut-off and c_n is the weighted sum of deprivations, given by:

$$c_n^t = \sum_{d=1}^{D} w_d I(x_{nd}^t \geq z_d) \tag{2}$$

where w_d is the weight of dimension $d = 1, \ldots, D$, x_{nd} is each individual's achievement for dimension d and z_d is the dimension specific cut-off. Finally, $I()$ is an indicator that takes the value of 1 if the expression in parenthesis is true, otherwise it takes the value of 0. Apart from the simple headcount ratio, this method offers the opportunity to calculate the intensity of deprivation suffered by poor people. In formal terms:

$$A(t) = \frac{1}{N^t H(t) D} \sum_{n=1}^{N^t} I(c_n^t \geq k) c_n^t \tag{3}$$

The combination of (1) and (3) gives the adjusted headcount ratio for multidimensional poverty (see (4)), which quantifies the weighted average number of deprivations. As it is shown, the adjusted headcount ratio is the product of the headcount ratio and intensity and thus is sensitive to changes in both measures.

$$M^0(t) = \frac{1}{N^t D} \sum_{n=1}^{N^t} I(c_n^t \geq k) c_n^t = H(t)A(t) \tag{4}$$

The fact that MPI is an adjusted headcount ratio gives researchers the opportunity to decompose this measure in numerous ways. First, it is possible to calculate each dimension's contribution to overall poverty (dimensional breakdown) and thus, specify the composition of multidimensional poverty. Defining the censored headcount ratio for each dimension (CHd) as the probability of being multidimensionally poor and at the same time deprived in dimension d (see (5)), it becomes clear that the adjusted headcount ratio will be a weighted sum of the censored headcount ratios. Using (6) the contribution of each dimension to the overall MPI can then be calculated as $\frac{w_d}{D}\frac{\mathrm{CH}_d}{M_0}$.

$$\mathrm{CH}_d(t) = \frac{1}{N} \sum_{n=1}^{N} I(c_n^t \geq k \cap x_{nd}^t \leq z_d) \tag{5}$$

$$M^0(t) = \sum_{d=1}^{D} \frac{w_d}{D} \mathrm{CH}_d(t), \qquad \text{where} \quad \sum_{d=1}^{D} w_d = 1 \tag{6}$$

A non-overlapping subgroup decomposition can also be implemented on the available data in order for the underlying subgroup dynamics to be revealed. Using (7) it is possible to examine each group's $i = 1, \ldots, G$ contribution to the overall regional MPI:

$$M^0(t) = \sum_{i=1}^{G} \varphi_i^t M_i^0 \tag{7}$$

where $\varphi_i^t = (N_i^t)/N^t$ is the population weight of group i in period t.

Following the EU's efforts to adjust the multidimensional measurement of poverty framework under the concept of developed countries (Atkinson, Marlier 2010), the work of Weziak-Bialowolska, Dijkstra (2014) adjusts the main dimensions of MPI at a regional level for the case of EU regions. This includes the definition of the variables that will be incorporated in each dimension, as well as the corresponding weights and cut-offs, to fit a developed country framework.

Table 2 presents the core MPI dimensions by the EU, alongside their corresponding weights and cut-offs (Weziak-Bialowolska, Dijkstra 2014, p. 22). Given the fact that our research does not focus on the development of a novel methodology regarding the MPI calculation, we have used the conceptualisation, including weights and cut-offs, proposed by Weziak-Bialowolska, Dijkstra (2014). In their case, the weighting scheme does not imply equal weights between the three key dimensions, whereas equal weights are considered for the sub-dimensional components. All variables used to calculate the MPI values are included in the European Survey on Income and Living Conditions (EU-SILC) database, and cut-offs for the three deprivation indices in the case of the living standards dimension are defined by the EU-SILC methodology. Detailed definitions for all the variables used in this study are given in Appendix A.

As a next step, a cluster analysis was conducted, in order to classify municipalities into groups based on their multidimensional poverty level. First, Ward's hierarchical grouping method was applied, using the minimum variance criterion as the base for the objective function optimization (Ward 1963). According to this criterion the objective function value that needs to be minimized is the sum of the squared deviations from the group mean (ESS) and it is given by:

$$\mathrm{ESS} = \sum_{i=1}^{n} x_i^2 - \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right)^2 \tag{8}$$

Table 2: MPI dimensions and indicators, including their weights and cut-offs (Weziak-Bialowolska, Dijkstra 2014)

MPI Dimensions (cut-off: **2 out of 3**)
• **Health**, weight: **2/6**, cut-off: **2 out of 3**
– General Health, weight: 1/3
– Unmet medical need due to lack of affordability and accessibility, weight: 1/3
– Unmet dental need due to lack of affordability and accessibility, weight: 1/3
• **Education**, weight: **1/6**, cut-off: **1 out of 1**
– No educational attainment
• **Living Standards**, weight: **3/6**, cut-off: **1 out of 3**
– **Material Deprivation**, weight: **1/3**, cut-off: **3 out of 9**
* Material Deprivation Index — MDI
– **Housing Deprivation**, weight: **1/3**, cut-off: **2 out of 4**
* Multidimensional Poverty in Housing Index – MPHoI
– **Environment Deprivation**, weight: **1/3**, cut-off: **2 out of 3**
* Multidimensional Poverty in Environment Index – MPEnI

Notes: For an extended description of the weights and cut-offs please see Weziak-Bialowolska, Dijkstra (2014)

where x_i is the score of i-th individual, which in our case takes the municipal MPI values corresponding to the year 2006. To perform a valid comparison between the obtained cluster groups for the years 2006 and 2011, we do not perform an additional cluster analysis for the year 2011. Results coming from this hierarchical procedure are presented in the form of a dendrogram (Figure 1). As is shown, several discrete clusters can be derived within the metropolitan area of Athens, based on their MPI values for 2006. Specifically, the case of $k = 3$ has been highlighted on the dendrogram, pointing out the three discrete groups of municipalities that can be derived in this case.

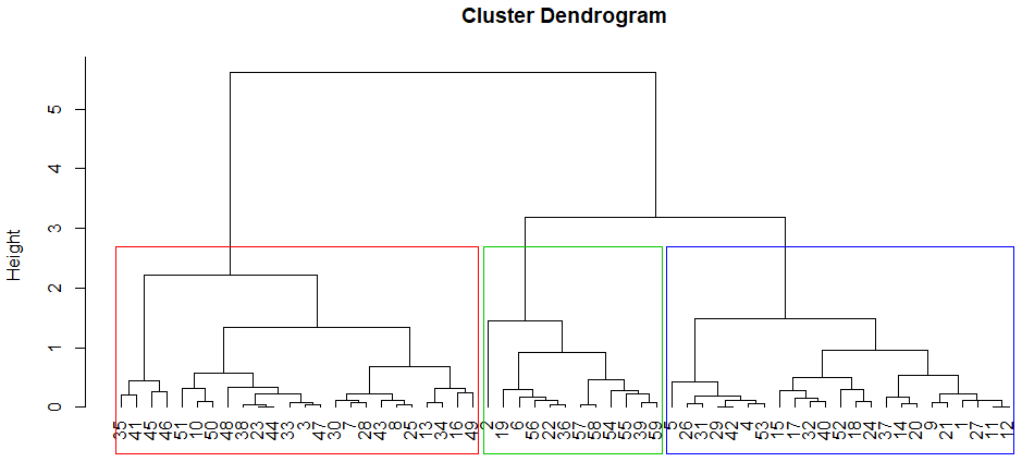
For further strengthening the robustness of the clustering process, k-means clustering was also applied. Being a non-hierarchical technique, k-means is based on partitioning a set of n observations $(x = x_1, x_2, \ldots, x_n)$ into k sets $(S = S_1, S_2, \ldots, S_k)$, and then minimize the within-cluster sum of squares (WCSS), or in other words the within cluster variance, as given below:

$$\text{WCSS} = \sum_{i=1}^{k} \sum_{x \in S_i} \|x - \mu_i\|^2 = \sum_{i=1}^{k} |S_i| \text{Var} S_i \tag{9}$$

Results coming from the k-means process are presented in the form of scatter plots (Figures 2 – 4). Three discrete diagrams are illustrated, covering the cases where $k = 3, 4$ or 5, respectively. These cases have been chosen based on the previous dendrogram results, that included all possible cluster combinations.

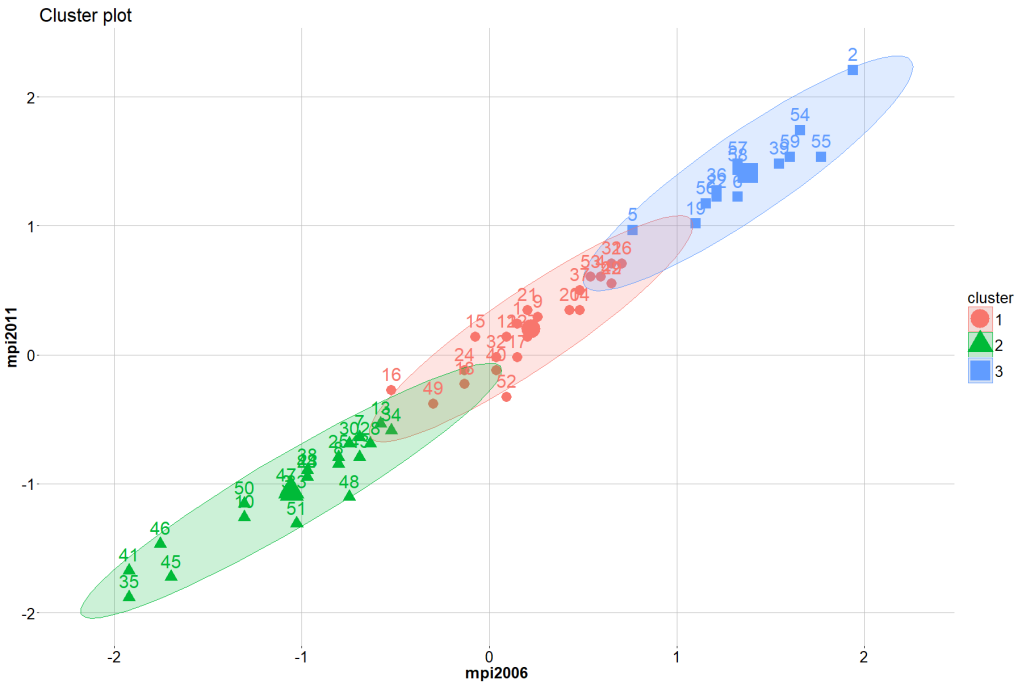
Moving on, the elbow method was used as the main criterion (Sugar 1998, Sugar, James 2003) to define the optimal number of clusters. According to it, the optimal number is specified with regards to the percentage of variance explained by the clusters, against the number of clusters. This means that the optimal number is reached, when the marginal gain of information of an additional cluster starts to decrease. Figures 5 and 6 illustrate the corresponding plots referring to the elbow criterion, for the cases of hierarchical and k-means clustering. In both cases, there is an evident "elbow" for $k = 3$, indicating this as the optimal choice for the number of clusters.

Finally, a cluster evaluation is needed to support the robustness of the previous results. In this case the stability of the cluster is assessed using the bootstrap method, for $B = 100$ resampling runs and $k = 3$ clusters. The results for each cluster are illustrated in Table



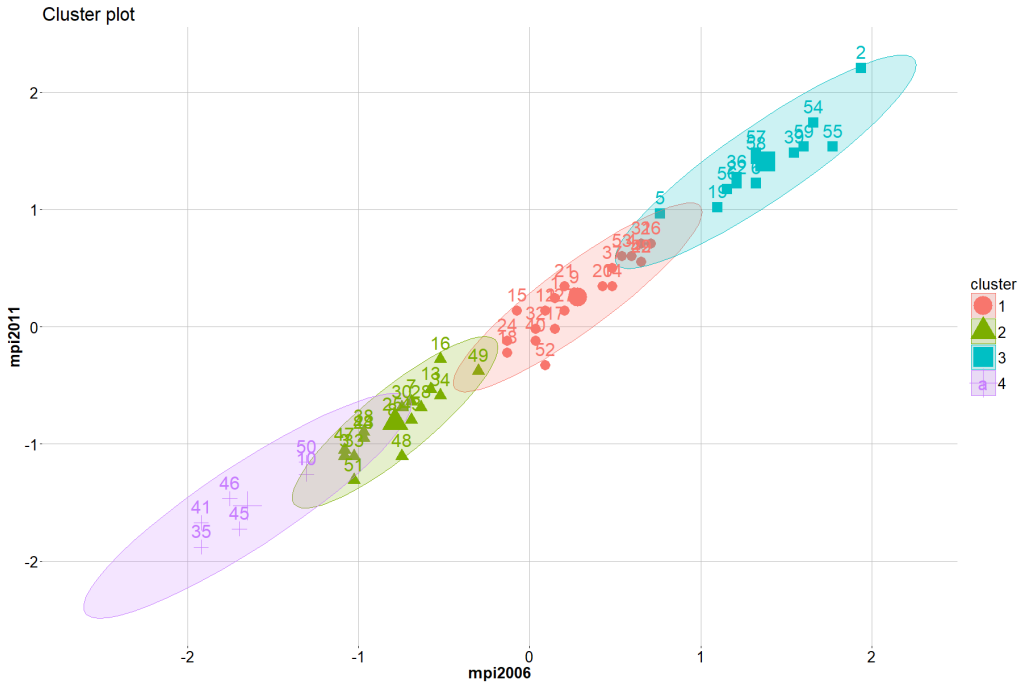
Source: Author's calculation
Notes: The numbers correspond to specific municipality IDs, which are given in Appendix B

Figure 1: Dendrogram for Ward's hierarchical clustering method based on the 2006 MPI values for municipalities in Athens



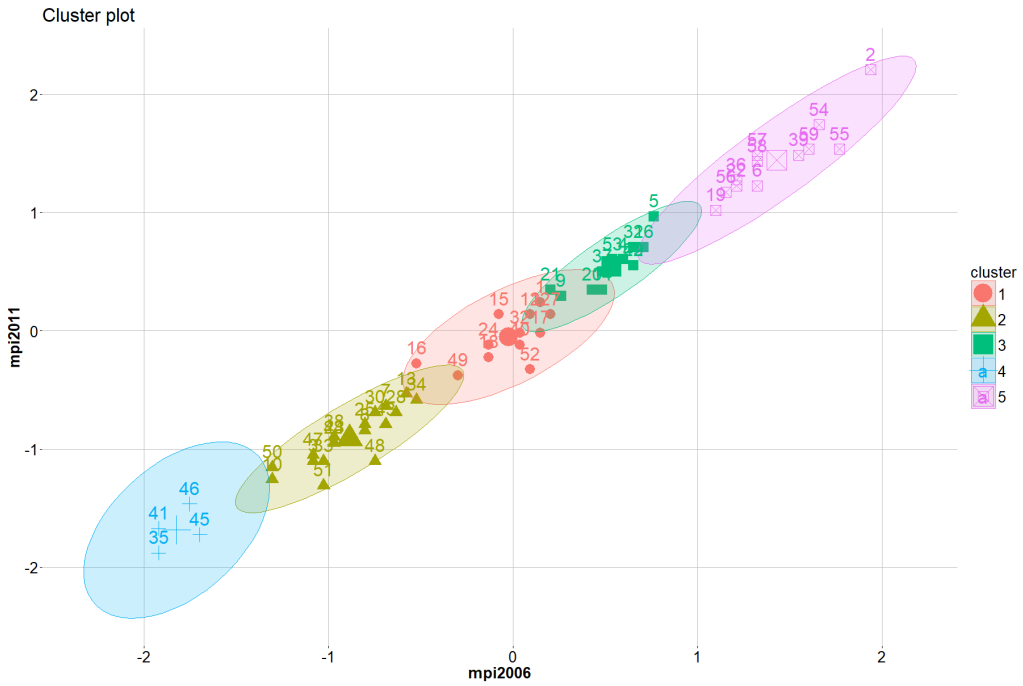
Source: Author's calculation
Notes: The numbers correspond to specific municipality IDs, which are given in Appendix B

Figure 2: Scatter plots for k-means clusters' solutions – k = 3



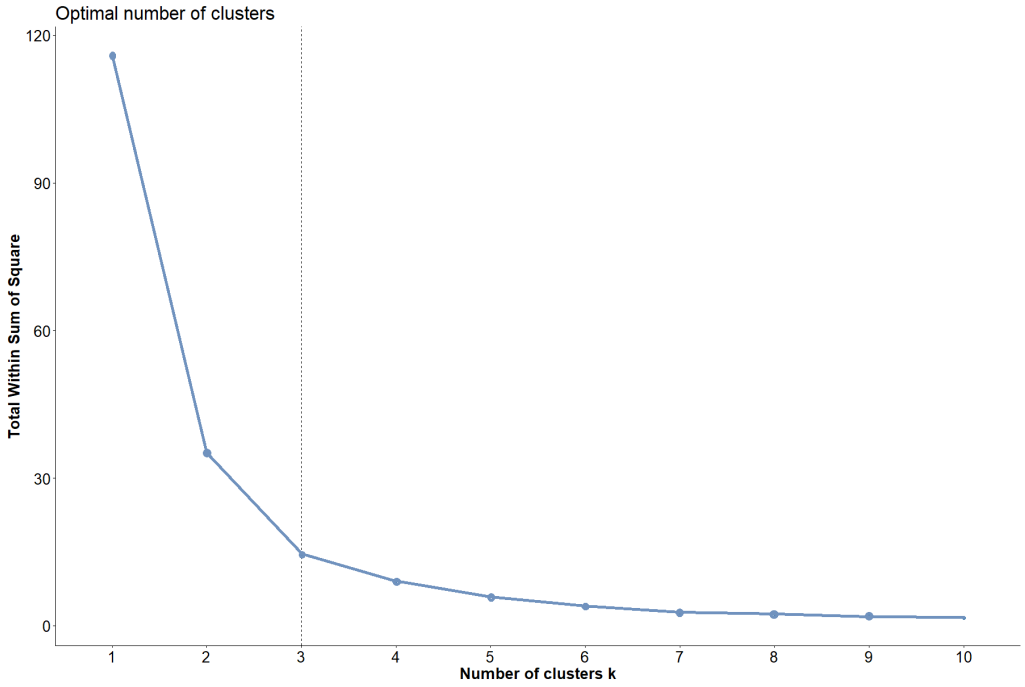
Source: Author's calculation
Notes: The numbers correspond to specific municipality IDs, which are given in Appendix B

Figure 3: Scatter plots for k-means clusters' solutions – k = 4



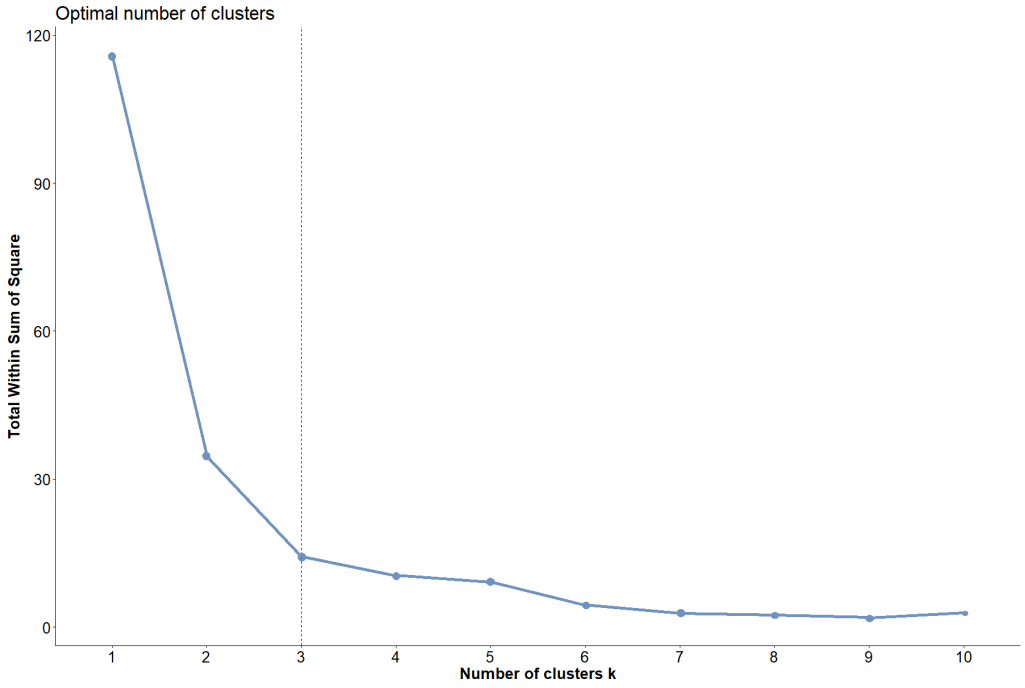
Source: Author's calculation
Notes: The numbers correspond to specific municipality IDs, which are given in Appendix B

Figure 4: Scatter plots for k-means clusters' solutions – k = 5



Source: Author's calculation

Figure 5: Identification of the optimal number of clusters using the elbow criterion for hierarchical clustering



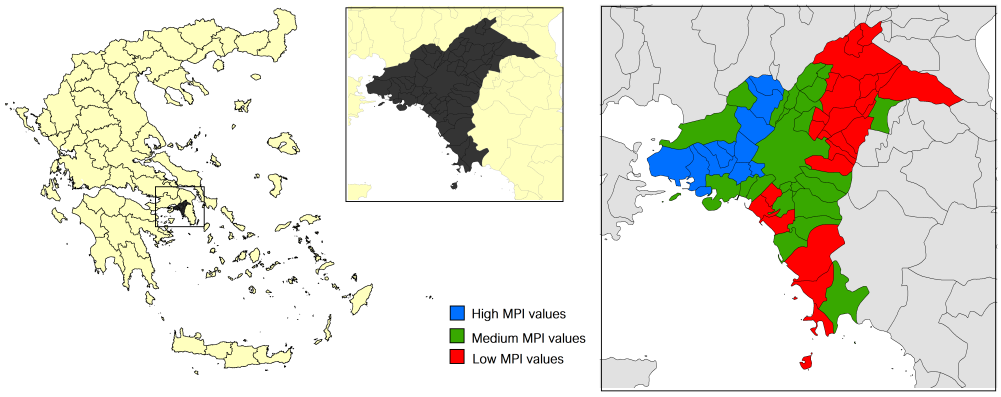
Source: Author's calculation

Figure 6: Identification of the optimal number of clusters using the elbow criterion for k-means clustering

Table 3: Results for bootstrap method of resampling

Cluster number	Jaccard bootstrap mean	Times a cluster has been dissolved	Times a cluster has been successfully recovered
1	0.906	0	91
2	0.923	1	88
3	0.947	0	99

Source: Author's calculation



Source: Panori et al. (2017) and author's calculations

Figure 7: Map illustrating three main clusters arising within the metropolitan area of Athens based on MPI values (2006)

3, including the Jaccard bootstrap mean value, the number of times a cluster has been dissolved and successfully recovered. In general, a valid, stable cluster should yield a mean Jaccard similarity value of 0.75 or more, whilst "highly stable" clusters should yield values of 0.85 and above (Hennig 2007, 2008). The values in our case indicate a high stability of the cluster results, obtained through the previous analysis.

This section presented a detailed description of the methodological framework used in this paper, regarding the calculation of the MPI index, under an EU country framework, as well as the process for building the clusters within the metropolitan area of Athens. It has also illustrated an assessment framework for the obtained clusters, in order to evaluate their stability.

5 Results

The aim of this paper is to reveal and compare the high, medium and low MPI clusters within the metropolitan area of Athens, in order to investigate for any existing structural differences of poverty between them. Figure 7 illustrates a spatial representation of the three derived clusters within Athens, based on the results of the previous analysis. A first observation is that the clusters illustrate high level of spatial concentration, indicating a social segregation pattern within Athens.

Some general characteristics for each cluster are shown in Table 4, focusing on further exploring this segregation pattern. By looking at the map, it becomes evident that high MPI municipalities span across the western part of Athens, encompassing highly deprived areas characterized by low mean income values. At the same time, the cluster that includes medium MPI municipalities is the largest one, in terms of population share, covering the central part of Athens, where the urban core lies. It is evident that municipalities belonging to this cluster are mostly medium-income areas. Finally, low MPI areas are located mainly in the north-east and south-east parts of Athens, encompassing

Table 4: Main characteristics of three clusters

Group	No. of obs	Pop. share 2006	Pop. share 2011	Mean income 06	Mean income 11	MPI 2006	MPI 2011
1/ High MPI	13	21.1	21.2	13177	11796	0.097	0.102
2/ Medium MPI	24	56.6	55.8	14297	12819	0.078	0.082
3/ Low MPI	22	22.3	23.0	15744	14175	0.059	0.060

Source: Author's calculations
Notes: Mean income is measured in €/capita.

Table 5: Calculated dimensional specific MPI values for each group

Group	2006			2011		
	Health	Education	Standard of Living	Health	Education	Standard of Living
1/ High MPI	0.041	0.491	0.109	0.059	0.462	0.126
2/ Medium MPI	0.038	0.359	0.108	0.056	0.347	0.120
3/ Low MPI	0.033	0.235	0.107	0.051	0.239	0.110

Source: Author's calculations

all high-income areas. Moreover, it is essential to notice that during the period under investigation there is no significant change in the population share of each group, a fact which helps to minimize the effects of population changes on the MPI evolution.

By decomposing the regional MPI index in its three main dimensions, additional differences arise between the clusters (Table 5). As expected, dimension-specific values remain larger for the high MPI cluster, in all cases. Moreover, it is interesting to notice the significant variation between high and low MPI clusters in the case of education, as well as the diversified reaction of this dimension during the period 2006-2011. More specifically, education illustrates a positive performance throughout this period, despite the deterioration in health and living conditions. This is a finding that could be used as preliminary evidence, in favour of claiming that during periods of economic recession, education may be treated as an investment opportunity, especially by the low-income classes.

Moving on, Table 6 presents the cluster headcount ratios for the three key dimensions. First, low headcount ratios for health (4-6% for 2006 and 7-8% for 2011) indicate a relatively small percentage of people being deprived in terms of this dimension. On the other hand, educational headcount ratios indicate a significant gap between the high and low MPI clusters. The share of educationally deprived people, located in a municipality that belongs to cluster 1, is more than two times higher when compared to the share of educationally deprived residents of cluster 3. Another finding is the increase in the living conditions' headcount ratios, which despite its very low level, is not similar in all groups. More specifically, the less advantaged group (cluster 1) was affected the most.

Table 7 presents the contribution of each dimension to the overall regional MPI value,

Table 6: Headcount ratio (%) for dimensional specific MPI values for each group

Group	2006			2011		
	Health	Education	Standard of Living	Health	Education	Standard of Living
1/ High MPI	5.7	49.1	24.8	8.2	46.2	27.9
2/ Medium MPI	5.2	35.9	24.8	7.7	34.7	26.7
3/ Low MPI	4.6	23.5	24.6	7.0	23.9	24.6

Source: Author's calculations

Table 7: Contribution (%) of each dimension to the overall MPI within each group

Group	2006			2011		
	Health	Education	Standard of Living	Health	Education	Standard of Living
1/ High MPI	23.21	35.35	41.44	24.65	34.36	40.99
2/ Medium MPI	24.96	33.49	41.55	27.34	32.18	40.48
3/ Low MPI	27.55	30.07	42.38	29.91	29.92	40.17

Source: Author's calculations

Table 8: Contribution (%) of each group to the overall regional MPI

Group	2006			2011		
	Contribution (a)	Population share (b)	Ratio (a)/(b)	Contribution (a)	Population share (b)	Ratio (a)/(b)
1/ High MPI	26.3	21.1	1.25	26.6	21.2	1.26
2/ Medium MPI	56.8	56.6	1.01	56.4	55.8	1.01
3/ Low MPI	16.9	22.3	0.76	17.0	23.0	0.74

Source: Author's calculations

in order to provide some additional information regarding the structural characteristics of multidimensional poverty within each cluster. It should be noted that there are no great differences between the three clusters, which illustrate a similar structure of multidimensional poverty through space. In general, there is an increased participation of the living conditions component in all cases, whilst the educational dimension is the second most important component of multidimensional poverty, especially in the case of low-income areas (cluster 1).

Besides investigating the structure of multidimensional poverty in each cluster separately, it is also very interesting to examine the results that arise when combining all data in order to represent the contribution of each cluster to the overall regional MPI of the metropolitan area of Athens. Table 8 illustrates the results based on (7). As expected, the main contribution to the overall MPI comes from the medium MPI cluster, which is the most populated, including the urban core of the city. Nonetheless, it is important to point out that there is a noticeable difference in the case of the other two (high and low MPI) clusters, when comparing their population shares with the share of their relative contribution to the overall regional MPI.

Although the high MPI cluster has a population share of 21.20% in 2011, its contribution to the overall regional MPI reaches 26.64%. This is expected, as it includes highly deprived areas, which contribute to a larger extent to the overall poverty levels within Athens. On the other hand, the population share of the low MPI cluster is 23%, whereas its contribution on multidimensional poverty is 17.00%. This difference becomes clear when looking directly at the ratio columns. Values higher than 1 indicate a contribution to the regional MPI which is higher than the corresponding population share, and vice versa. For the case of the medium MPI cluster, the ratio is almost equal to unity, indicating a balanced behaviour due to its large population share.

In order to explore any existing differences regarding the dynamic behaviour of the three clusters, it is essential to show the absolute and relative changes of the MPI values within each group (Table 9). Starting from the overall regional MPI index, both absolute and relative changes increase as we move to higher MPI clusters. A similar trend is also observed in the case of the standard of living component. In this case, the high MPI cluster experiences a sharp relative increase between 2006 and 2011, which is almost 4.5 times higher (15.95%/3.60%=4.43) when compared to the corresponding change in the low MPI cluster. Nonetheless, these variations seem to have different behaviour when it comes to health and education components. In the first case, an opposite movement is observed, when moving from cluster 1 to cluster 3, which does not seem to be in line with the corresponding absolute changes, due to differences between the initial values.

Table 9: Absolute and relative changes 2006/11 of the MPI values within sub-groups and for each dimension separately

		Total	High MPI	Medium MPI	Low MPI
Regional MPI	Absolute	0,003	0,004	0,003	0,001
	Relative (%)	3,55	4,38	3,93	1,82
Health	Absolute	0,018	0,018	0,018	0,017
	Relative (%)	47,21	43,00	47,62	51,36
Education	Absolute	-0,013	-0,029	-0,012	0,003
	Relative (%)	-3,54	-5,83	-3,39	1,39
Standard of Living	Absolute	0,012	0,017	0,012	0,004
	Relative (%)	10,68	15,95	11,54	3,60

Source: Author's calculations

However, in the case of education, a decrease on MPI values for high and medium MPI clusters is observed, whilst the low MPI cluster seems to experience a relative increase of 1.39% on its educational MPI component.

The previous analysis has shown that only small variations exist between the derived clusters. Any distributional and/or behavioural differences reflects the underlying urban dynamics within the metropolitan area of Athens. These encompass complex processes, such as economic restructuring, migration and the welfare state, leading to different socio-spatial outcomes. Economic disturbances, including the 2008 economic crisis, constitute additional forces, that push poverty structures, alongside with their dimensional components, into similar or opposite directions. The evolution of education and living standards are two indicative examples, regarding the diversified impact of urban dynamics on these two dimensions of poverty.

6 Concluding remarks

This paper focuses on exploring poverty concentration and structure within the metropolitan area of Athens. The study tried to investigate poverty under a multidimensional framework, revealing the necessity of expanding the notion of that kind of urban phenomena to other non-income-based dimensions. Using the MPI approach to identify and map poverty concentration patterns within the city of Athens, has yielded very interesting findings.

First, a cluster analysis has been performed to define the optimal number of clusters. The results, coming from both hierarchical and partitioning techniques, have shown that the division of the metropolitan area of Athens into three discrete groups of municipalities, based on their MPI values for the year 2006, is the best way to approach multidimensional poverty segregation. The analysis revealed the existence of a spatial concentration of multidimensionally poor areas on the west suburban area of Athens, being traditionally inhabited by low-income workers. At the same time, less deprived municipalities of the city are primarily located on the south-east and north-east suburbs of Athens, whilst the urban core of the city seems to be characterised by medium-level MPI values. The existing differences between these clusters, in both income and MPI values, indicated the necessity of expanding the analysis, in order to explore any structural variances.

When looking at the structural and dynamic features within separate clusters, there are no considerable differences in terms of health and living conditions. An increase in health values similar for all groups has been noticed, whereas living conditions seem to have deteriorated slightly more in already deprived areas. More specifically, areas being part of the high MPI cluster seem to have been mostly affected in terms of material deprivation, housing problems and environment. Given the fact that during the period under investigation extended recessionary economic transitions have been taking place within Greece, this finding is in line with Andersson, Hedman (2016), which point out that economic crisis exhibits geographically uneven outcomes affecting poor areas the most.

Furthermore, it is essential to point out an existing significant gap in the educational component. As expected, highly deprived suburban areas show much lower educational attainment. However, there seems to be a convergence of the two tails of the educational distribution during the period 2006-2011, as the gap between high and low MPI areas becomes smaller. Given the high complexity of the social processes taking place at the urban level, more research should be made towards identifying the reasons for these movements. Possible reasons for this behaviour might include first, a high level of positive social mobility in those areas, not being followed by an additional residential mobility, and second, the brain drain phenomenon that has broadly affected high-income areas.

Existing variations in between-group behaviour highlight the necessity for expanding urban poverty research, especially in cases of large metropolitan areas. Through this paper, a step has been made towards a better understanding of the socio-spatial dialectic within the city of Athens, aiming to reveal the underlying interactions that form and maintain various aspects of poverty. Although it is difficult to generalise the conclusions of this specific case study, the implementation of similar analyses to other urban areas could offer a more holistic view regarding the structure of poverty within different socio-economic contexts.

Acknowledgement

I would like to thank the editor and two anonymous referees for their insightful comments.

References

Alkire S, Foster J (2011a) Counting and multidimensional poverty measurement. *Journal of Public Economics* 95[7-8]: 476–487. CrossRef.

Alkire S, Foster J (2011b) Understandings and misunderstandings of multidimensional poverty measurement. *The Journal of Economic Inequality* 9[2]: 289–314. CrossRef.

Allard S, Paisner S (2016) The rise of suburban poverty. Oxford handbooks online. Retrieved 11 nov. 2016, from http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199935307.001.0001/oxfordhb-9780199935307-e-96. CrossRef.

Andersson R, Hedman L (2016) Economic decline and residential segregation: A Swedish study with focus on Malmö. *Urban Geography* 37[5]: 748–768. CrossRef.

Apablaza M, Yalonetzky G (2013) Decomposing multidimensional poverty dynamics. Young lives working paper 101, https://ora.ox.ac.uk/objects/uuid:cde591e7-fd51-4d1b-a4ee-fe77f76ce5df/datastreams/ATTACHMENT01

Atkinson B, Marlier E (2010) Income and living conditions in Europe. European Union, Luxembourg, http://ec.europa.eu/eurostat/documents/3217494/5722557/KS-31-10-555-EN.PDF/e8c0a679-be01-461c-a08b-7eb08a272767

Berube A, Frey WH (2002) A decade of mixed blessings: Urban and suburban poverty in census 2000. Brookings Institution, Center on Urban and Metropolitan Policy, Washington, DC

Bischoff K, Reardon SF (2013) Residential segregation by income, 1970-2009. Brown University, https://s4.ad.brown.edu/Projects/Diversity/Data/Report/report10162013.pdf

Cooke T, Marchant S (2006) The changing intrametropolitan location of high-poverty neighbourhoods in the US, 1990-2000. *Urban Studies* 43[11]: 1971–1989. CrossRef.

Cooke TJ (2010) Residential mobility of the poor and the growth of poverty in inner-ring suburbs. *Urban Geography* 31[2]: 179–193. CrossRef.

Cooke TJ, Denton C (2015) The suburbanization of poverty? An alternative perspective. *Urban Geography* 36[2]: 300–313. CrossRef.

Duncan OD, Duncan B (1955) Residential distribution and occupational stratification. *American journal of sociology* 60[5]: 493–503. CrossRef.

Gotham KF (2003) Toward an understanding of the spatiality of urban poverty: The urban poor as spatial actors. *International Journal of Urban and Regional Research* 27[3]: 723–737. CrossRef.

Green AE (1998) The geography of earnings and incomes in the 1990s: An overview. *Environment and Planning C: Government and Policy* 16[6]: 633–647. CrossRef.

Hamnett C (2003a) Gentrification and the middle-class remaking of inner London, 1961–2001. *Urban studies* 40[12]: 2401–2426. CrossRef.

Hamnett C (2003b) *Unequal City: London in the Global Arena*. Routledge, Milton Park. CrossRef.

Harris R, Longley P (2004) Targeting clusters of deprivation within cities. In: Stillwell J, Clarke G (eds), *Applied GIS and spatial analysis*. Wiley, Chichester, 89–110. CrossRef.

Hennig C (2007) Cluster-wise assessment of cluster stability. *Computational Statistics and Data Analysis* 52: 258–271. CrossRef.

Hennig C (2008) Dissolution point and isolation robustness: Robustness criteria for general cluster analysis methods. *Journal of Multivariate Analysis* 99: 1154–1176. CrossRef.

Hunter P (2014) Poverty in suburbia: A Smith Institute study into the growth of poverty in the suburbs of England and Wales. The Smith Institute, https://smithinstitute-thinktank.files.wordpress.com/2014/09/poverty-in-suburbia.pdf

Jargowsky P (2003) Stunning progress, hidden problems: The dramatic decline of concentrated poverty in the 1990s. Brookings institution, Washington, DC

Jargowsky PA (1996) Beyond the street corner: The hidden diversity of high-poverty neighborhoods. *Urban Geography* 17[7]: 579–603. CrossRef.

Jargowsky PA (2002) Sprawl, concentration of poverty, and urban inequality. In: Squires G (ed), *Urban sprawl: Causes, consequences, and policy responses, 39-72*. The Urban Institute Press, Washington, DC. CrossRef.

Jargowsky PA (2013) Concentration of poverty in the new millennium: Changes in the prevalence, composition, and location of high-poverty neighborhoods. Century foundation, New York

Kalogirou S (2011) Spatial inequalities and interpretative factors for the geographical distribution of declared income in Greece. *Aihoros* 11: 68–101

Kingsley GT, Pettit KL (2003) Concentrated poverty: A change in course. Urban institute, http://www.urban.org/sites/default/files/publication/58946/310790-Concentrated-Poverty-A-Change-in-Course.PDF

Kneebone E (2014) The growth and spread of concentrated poverty, 2000 to 2008–2012. Brookings Institution, https://www.brookings.edu/interactives/the-growth-and-spread-of-concentrated-poverty-2000-to-2008-2012/, Washington, DC

Kneebone E, Berube A (2013) Confronting suburban poverty in America. Brookings Institution, Washington, DC, https://www.brookings.edu/book/confronting-suburban-poverty-in-america/

Lee N, Adam D, Rodríguez-Pose A, Green A, Atfield G (2014) Cities, growth and poverty: Evidence review. Joseph Rowntree Foundation, York, UK

Maloutas T (2000) *The Social and Economic Atlas of Greece*, Volume I: The Cities National Centre for Social Research. EKKE-University of Thessaly Press, Athens/Volos

Maloutas T (2001) The social morphology of cities. In: Maloutas T (ed), *Social and Economic Atlas of Greece*, Volume 1: The Cities. EKKE-University of Thessaly Press, Athens/Volos

Maloutas T (2004) Segregation and residential mobility spatially entrapped social mobility and its impact on segregation in Athens. *European Urban and Regional Studies* 11[3]: 195–211

Maloutas T (2007) Segregation, social polarization and immigration in Athens during the 1990s: Theoretical expectations and contextual difference. *International Journal of Urban and Regional Research* 31[4]: 733–758. CrossRef.

Maloutas T (2015) Socio-economic segregation in Athens at the beginning of the twenty-first century. In: Tammaru T, van Ham M, Marcińczak S, Musterd S (eds), *Socio-Economic Segregation in European Capital Cities: East Meets West*. Routledge, Milton Park, 156–185

Maloutas T, Arapoglou V, Kandylis G, Sayas J (2012) Social polarization and de-segregation in Athens. In: Maloutas T, Fujita K (eds), *Residential Segregation in Comparative Perspective. Making Sense of Contextual Diversity*. Routledge, Milton Park, 257–285

Maloutas T, Emmanuel D, Pantelidou-Malouta M (2006) *Athens-Social Structures, Practices and Attitudes: New Parameters and Trends 1980–2000*. National Center for Social Research, Athens

Massey DS (2007) *Categorically Unequal: The American Stratification System*. Russel Sage Foundation, New York

Massey DS, Denton NA (1993) *American apartheid: Segregation and the making of the underclass*. Harvard University Press, Cambridge, Mass.

McMullen SC, Smith WJ (2002) The geography of poverty in the Atlanta region. Geogria State University, http://www.academia.edu/download/30771264/report7.pdf

Musterd S (2005) Social and ethnic segregation in Europe: Levels, causes, and effects. *Journal of urban affairs* 27[3]: 331–348. CrossRef.

Musterd S, Marcinczak S, van Ham M, Tammaru T (2015) Socio-economic segregation in European capital cities: Increasing separation between poor and rich. Forschungsinstitut zur Zukunft der Arbeit (IZA), IZA Discussion Paper 9603, http://ftp.iza.org/dp9603.pdf

Noble M, Wright G, Dibben C, Smith GAN, McLennan D, Anttila C, Barnes H, Mokhtar C, Noble S, Avenell D, Gardner J, Covizzi I, Lloyd M (2004) Indices of deprivation 2004. Report to the Office of the Deputy Prime Minister, Neighbourhood Renewal Unit, London

Noble M, Wright G, Smith G, Dibben C (2006) Measuring multiple deprivation at the small-area level. *Environment and planning A* 38[1]: 169–185. CrossRef.

Orford S (2004) Identifying and comparing changes in the spatial concentrations of urban poverty and affluence: A case study of inner London. *Computers, Environment and Urban Systems* 28[6]: 701–717. CrossRef.

Pacione M (2004) Environments of disadvantage: Geographies of persistent poverty in Glasgow. *Scottish Geographical Journal* 120[1-2]: 117–132. CrossRef.

Panori A, Ballas D, Psycharis Y (2017) SimAthens: A spatial microsimulation approach to the estimation and analysis of small area income distributions and poverty rates in the city of Athens, Greece. *Computers, Environment and Urban Systems* 63: 15–25. CrossRef.

Pantazis P, Psycharis Y (2016) Residential segregation based on taxable income in the metropolitan area of Athens. In: Maloutas T, Spyrellis S (eds), *Athens Social Atlas. Digital compendium of texts and visual material.* http://www.athenssocialatlas.gr/en/article/income-groups/

Park RE, Burgess EW, McKenzie RD (1925) *The City.* The University of Chicago Press, reprint 1984, Chicago

Pinch S (1993) Social polarization: a comparison of evidence from Britain and the United States. *Environment and Planning A* 25[6]: 779–795. CrossRef.

Reardon SF, Bischoff K (2011) Income inequality and income segregation. *American Journal of Sociology* 116[4]: 1092–1153. CrossRef.

Sassen S (1991) *The Global City: New York, London, Tokyo.* Princeton University Press, Princeton, NJ

Scott AJ, Storper M (2015) The nature of cities: The scope and limits of urban theory. *International Journal of Urban and Regional Research* 39[1]: 1–15. CrossRef.

Soja EW (1989) *Postmodern Geographies: The Reassertion of Space in Critical Social Theory.* Verso, Brooklyn, NY

Sugar CA (1998) Techniques for clustering and classification with applications to medical problems. Doctoral dissertation, Stanford University

Sugar CA, James GM (2003) Finding the number of clusters in a dataset: An information-theoretic approach. *Journal of the American Statistical Association* 98[463]: 750–763. CrossRef.

Ward JH (1963) Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58[301]: 236–244. CrossRef.

Watson T (2009) Inequality and the measurement of residential segregation by income in American neighborhoods. *Review of Income and Wealth* 55[3]: 820–844. CrossRef.

Weziak-Bialowolska D, Dijkstra L (2014) Monitoring multidimensional poverty in the regions of the European Union. Analysis of situation in 2012 and 2007. European Union, Joint Research Centre

A Appendix A

This appendix presents in detail all indicators used to calculate the MPI components, based on Weziak-Bialowolska, Dijkstra (2014) conceptualization of regional MPI for the EU countries. All variables are included in the EU-SILC database.

Dimension 1: Health (cut-off: 2 out of 3)

Component 1.1: General health:

PH010 — Reporting bad or very bad conditions.

Component 1.2: Unmet medical need due to lack of affordability and accessibility

PH040, PH050 — Unmet need for medical examination or treatment because it was not affordable, there was a waiting list or it was too far to travel/no means of transportation.

Component 1.3: Unmet dental need due to lack of affordability and accessibility

PH060, PH070 — Unmet need for dental examination or treatment because it was not affordable, there was a waiting list or it was too far to travel/no means of transportation.

Dimension 2: Education (cut-off: 1 out of 1)

Component 2.1: Educational attainment

PB010, PB140, PE010, PE040 — A person of more than 24 years not having at least upper secondary education or in the age range 16-24 years who has finished no more than lower secondary education and is not involved in further education (based on early school leaver definition).

Dimension 3: Living Standards (cut-off: 1 out of 3)

Component 3.1: Material Deprivation (3 out of 9)

HS070, HS090, HS100, HS110 — Household cannot afford: a telephone (including a mobile phone), a computer, a washing machine, a car.

HS010, HS011, HS020, HS021 — Households with arrears on mortgage or rent payments or utility bills.

HS040 — Lack of capacity in a household to afford paying for one-week annual holiday away from home.

HS050 — Lack of capacity in a household to afford a meal with meat, chicken, fish (or vegetarian equivalent) every second day.

HS060 — Lack of capacity to face unexpected financial expenses.

HH050 — Household without ability to keep home adequately warm.

Component 3.2: Housing problems (cut-off: 2 out of 4)

HH030 — Crowding index (average number of people per room available to the household) larger than 2

HH040, HH080/HH081, HS160 — Problems with dwelling: - leaking roof, damp walls/-floors/foundation, or rotten window frames or floor – too dark, not enough light – without bath or shower for sole use in dwelling.

Component 3.3: Environment (cut-off: 2 out of 3)

HS170, HS180, HS190 — Household experiences: - noise from neighbours or from the street – pollution, grime or other environmental problems – crime, violence or vandalism in the area.

B Appendix B

This appendix illustrates the list of municipalities within the metropolitan area of Athens, alongside with their ID numbers.

ID	Municipality	ID	Municipality
1	Athens	31	Nea Filadelfeia
2	Agia Varvara	32	Nea Chalkidona
3	Agia Paraskevi	33	Neo Psychiko
4	Agios Dimitrios	34	Palaio Faliro
5	Agioi Anargyroi	35	Papagou
6	Aigaleo	36	Peristeri
7	Alimos	37	Petroupoli
8	Amarousio	38	Pefki
9	Argyroupoli	39	Tavros
10	Vrilissia	40	Ymittos
11	Vyronas	41	Filothei
12	Galatsi	42	Chaidari
13	Glyfada	43	Chalandri
14	Dafni	44	Cholargos
15	Elliniko	45	Psychiko
16	Zografou	46	Ekali
17	Ilioupoli	47	Nea Penteli
18	Irakleio	48	Penteli
19	Ilio	49	Vari
20	Kaisariani	50	Voula
21	Kallithea	51	Vouliagmeni
22	Kamatero	52	Gerakas
23	Kifisia	53	Piraeus
24	Lykovrysi	54	Agios Ioannis Rentis
25	Melissia	55	Drapetsona
26	Metamorfosi	56	Keratsini
27	Moschato	57	Korydallos
28	Nea Erythraia	58	Nikaia
29	Nea Ionia	59	Perama
30	Nea Smyrni		

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 3, 2017, 39–63
DOI: 10.18335/region.v4i3.166

journal homepage: region.ersa.org

Effective Clusters as Territorial Performance Engines in a Regional Development Strategy – A Triple-Layer DEA Assessment of the Aviation Valley in Poland

Karima Kourtit[1], Peter Nijkamp[2], Soushi Suzuki[3]

[1] KTH Royal Institute of Technology, Stockholm, Sweden; Adam Mickiewicz University, Poznan, Poland (email: karima.kourtit@abe.kth.se)
[2] KTH Royal Institute of Technology, Stockholm, Sweden; Tinbergen Institute, Amsterdam, The Netherlands; Adam Mickiewicz University, Poznan, Poland (email: pnijkamp1@gmail.com)
[3] Hokkai-Gakuen University, Sapporo, Japan (email: soushi-s@lst.hokkai-s-u.ac.jp)

Received: 29 September 2016/Accepted: 18 May 2017

Abstract. In the past decades, a new concept has been introduced in the regional development literature, viz. economic-technological clusters. A wealth of studies has been published on the conceptual, operational and policy foundation, and relevance of this concept, especially in relation to previously developed regional growth concepts, such as industrial districts, industrial complexes, or growth centers. In the present paper, clusters will be regarded as the spatial foci of sustainable territorial performance strategies and synergetic actions by both public and private actors. This paper aims to address the relevance of cluster concepts for an effective regional development policy, based on the notion of territorial capital. It does so by introducing a new concept, viz. effective cluster, in which spatial-economic synergy, local/regional concentration of industry, and the supporting role of territorial capital are regarded as the main determinants of a highly performing cluster in a given territory. The effective cluster concept will be illustrated and tested on the basis of a field study on the aviation and aerospace cluster 'Dolina Lotnicza' in the Podkarpackie region in South-East Poland. This is one of the most vibrant high-tech clusters in the country. Rather than providing a critical assessment of the specific development strategy of the Podkarpackie region, this study will show the added value of a new general conceptual framework based on effective clusters. A novel approach based on a triple-layer architecture will be adopted here, viz. a quantitative comparative analysis of the 16 Polish 'voivodships' (main administrative regions in the country, at a NUTS-2 level), a benchmark analysis of the 25 counties ('powiats') within the Podkarpackie voivodship (at a NUTS-4 level), and an effective industrial cluster analysis on the basis of the individual aviation firms located in the Podkarpackie region. In each step, an extended Data Envelopment Analysis (DEA), characterized by a merger of a Slack-Based Measure (SMB) and a super-efficiency (SE) DEA model, will be used in order to achieve an unambiguous ranking of the various regions or relevant Decision Making Units (DMUs) in the area concerned. The study will employ an extensive database based on field work among the individual actors in the cluster, in combination with a broadly composed territorial-capital database for the areas under study. The paper will be concluded with some strategic policy lessons.

1 Introduction: Aviation as a Smart Specialization

Over the past decades, regional development policy has sought to remove the deficiencies in the economic structure of less developed regions through the development of effective infrastructures and the use of appropriate knowledge and innovation systems. Clearly, modern regional development policy seeks to exploit often the economic potential of promising – though not necessary lagging – regions through the creation of accessible transport and communication systems and the transmission of advanced knowledge among actors (see Acz et al. 2009). An important intervening opportunity for an effective use and dissemination of knowledge among actors is formed by social capital (see Westlund 2006). Against this background, regions in an advanced knowledge economy and an open systems' network may act as an economic and technological engine for accelerated knowledge-based spatial development.

Nowadays, there is an increasing awareness that regions should exploit their competitive advantage through a focus on those cognitive activities in which the regions concerned have a proven excellence. This has prompted the notion of smart specialization. The rationale of this concept is based on the idea that in an open world, regions and industries have to compete for the most efficient market strategy, so as to maximize their revenues through a specialization in the most productive or efficient activities. Consequently, regional clusters of industries should not engage in a broad and unfocussed portfolio of industrial or service activities, but should rather seek to optimize their market position on the basis of a smart choice of a limited number of specialized industries (see Batabyal, Nijkamp 2015), such as medical technology, nanotechnology, or environmental technology.

A good example of regional smart specialization is formed by the airline industry. Aviation is a rapidly growing industry world-wide. For example, it is sometimes argued that in the decades to come at least 30,000 new aircraft would have to be built so as to meet the rising demand of mobile people. In the light of this large-scale global development, this dynamic offers, of course, a great opportunity for regional development, provided a proper aviation specialization is strived for from a place-based perspective.

Central and Eastern Europe has a long-standing tradition in the development of aircraft and related products, in particular the Czech Republic, Romania, and Poland. After the fall of the iron curtain, these countries have aimed to employ the historical roots of their former strong aviation industry as an anchor point for developing new regional spearheads in the aerospace industry (see for a review Bochniarz 2007). One of the regions which has in recent years played an active role in the redevelopment of a modern aviation sector is the Podkarpackie region in the South-Eastern part of Poland. It has heavily invested in new infrastructures, advanced knowledge, and international linkages so as to build up a modern aerospace industry. To emphasize the smart aviation specialization in this region, it has in recent years been baptized as the Aviation Valley (see Kaszuba 2012).

The smart specialization in aerospace activities in a given region is based on the assumption that this knowledge-intensive and innovation-driven activity may form a cornerstone for new advanced activities in the area concerned. This modern cluster specialization may generate high revenues, many new jobs, and a high international profile. This calls of course for a thorough (mainly quantitative) assessment of all relevant effects of such an aviation cluster. A good example of such an impact study, on the Boeing aviation cluster in Washington State (USA), can be found in CAI (2013) and Sommers et al. (2008). Additionally, in recent years, various studies in Europe have been carried out on the (regional-) economic significance of the aviation sector. Some interesting examples can be found in a study for the European Commission (ECORYS 2009), and in a study on aviation networks (Zuliani, Jalebert 2005). A policy-oriented study on the regional aerospace cluster policy in Europe can be found in Schönfeld, Jouaillec (2008).

It goes without saying that in the light of the economic and cognitive importance of high-quality regional aviation initiatives, proper decisions have to be taken on the choice of the portfolio of activities, on the skill levels needed, on the necessary infrastructural provisions, and on the set of innovation strategies needed to ensure the highest performance of the cluster concerned.

In the light of the previous observations, this paper aims to assess the success factors of the Polish Aviation Valley by seeking to arrive at a relative performance ranking of key actors, by using a 3-stage Data Envelopment Analysis (DEA), for an effective cluster constellation to be evaluated in three successive steps: (i) a comparative evaluation of the performance of the Podkarpackie region in South-East Poland vis-à-vis the other main administrative regions ('voivodships'); (ii) a comparative evaluation of the performance of the 25 counties ('powiats') in the Podkarpackie region that are the home base of the various firms in the Aviation Valley in Podkarpackie; and (iii) a comparative evaluation of the micro performance of the most important aviation business firms in this cluster on the basis of detailed interview information from these firms. The methodological toolbox for this triple-layer assessment comprises in particular: principal component analysis (PCA) and Data Envelopment Analysis (DEA).

The present study is organized as follows. After this introductory section, Sections 2 and 3 will, respectively, be devoted to a conceptual exploration and a concise description of the economic backgrounds and the current cluster position of the above mentioned Aviation Valley in Poland. In Section 4, we will present the methodological backgrounds of our territorial performance model, followed by a description of the database collected and employed for our research purposes. In Section 5, the operational model will be treated in the form of a comparative benchmark analysis of the constituents of the triple-layer architecture, where territorial capital data will be used as input for the DEA. The subsequent section (Section 6) will be devoted to a description and interpretation of our results, while the final section will offer concluding remarks, in particular on policy lessons for the Aviation Valley.

2 Territories and Production Units in Clusters

Regional development is not 'manna from heaven', but the result of deliberate smart strategic choices, decisions, and actions of stakeholders in a given area. It is based on an effective policy effort – both public and private – to shape attractive conditions for accelerated sustainable growth in a geographically concentrated area (see Andersson et al. 2016). The achievement of such a goal needs the fulfilment of various prerequisites: economies of density of activities, multi-tasking synergy of smart development strategies, and existence and exploitation of internal and external network liaisons among all relevant actors (see for an overview Capello, Nijkamp 2009).

In the rich history of regional development policy a variety of instrumental concepts has come to the fore: industrial districts, growth poles, growth centers, geographical clusters, development axes, industrial complexes, special economic zones, high-tech parks, etc. Despite a diversity in meaning and scope of such spatial growth concepts, they all served to enhance the economic performance of the area concerned, with a view to formulating and implementing solid strategies in order to improve its socio-economic competitive profile and to reduce spatial disparities in a regional or (inter)national arena through various forms of scale and agglomeration advantages (Nijkamp 2016).

Countries, regions, and municipalities all over the world are exposed to the challenge and opportunity to improve continuously their position and to optimize their socio-economic performance, by economizing on the use of critical (internal and external) resources to enter a more promising – though also more risky – competitive spatial-economic environment in a globalizing system. Such a performance may comprise: welfare, socio-economic well-being, income, safety, employment, economic growth, (inter)national recognition, social cohesion, sustainable development, and so forth. It should be noted here that there is in general no unambiguous and measurable performance measure for such territories, in contrast to industrial organizations, where nowadays KPIs (key performance indicators) act as signposts for company strategies. Clearly, some attempts at designing such systematic information have been made in recent years. For example, in an urban context, the notion of XXQ (maximum quality of a city) has recently been advocated as an overarching policy objective including in particular economic, social, and ecological performance indicators of a city (see Nijkamp 2008, Kourtit 2014, 2015).

It seems plausible that any territory in the form of an interlinked spatial entity

(country, region, municipality) seeks to optimize its relative position in a balanced way in order to achieve the highest possible quality of welfare, living, and working in this area (see Nijkamp 2008, Kourtit 2014, 2015), as compared to others. These 'others' may be the direct neighbors in the vicinity, but they may also be found at the international stage (see for instance, the competition between global world cities such as New York, London, Paris, or Tokyo) (we refer here, for example, to Kourtit et al. 2013, Arribas-Bel et al. 2013). Regions in our world may be assumed to maximize their contribution to – and share in – the total 'performance production' at a global scale, or at least within a given relevant continent or country. This performance[1] may be measured on a multidimensional and even multilevel scale – including e.g., GDP per capita, employment, public facilities, socio-economic equality, and ecological quality as well as critical network elements – as part of a so-called 'Territorial Performance Index'(TPI), which may act as crucial location factors and drivers for various actors across geographical units (see also Camagni 2009). Such a TPI is clearly a latent variable and may be the result of either competitive strategies or cooperation initiatives, or both, with regard to players in the same domain (see also Healy, Cote 2001). In this context, Kaasa, Part (2008) argued that "an individual's achievements would be higher, if he or she competed and cooperated with others through different networks and common value systems" (p. 5). This also holds for territories (countries, regions, cities, etc.) which may be seen as multi-tasking production units, often in liaison with other territories.

In the light of the previous observations, we will assume in the present paper that each territory has a TPI which can be produced or created through a smart combination of various productive inputs, in particular, Human Capital (HC), Infrastructural Capital (IC) and Social Capital (SC) (see also Becker 2009, De la Fuente, Vives 1995, Heckman 2000, Ravikumar, Glomm 1992, Rodriguez-Pose, Fratesi 2004). The TPI production function is thus shaped through at least three distinct – but often mutually connected – production factors. The welfare of regions is thus critically dependent on its resource use (Nijkamp 2016). The blend of these three categories of production inputs will be coined here 'Territorial Capital' (TC).

'Territorial Capital' is a new concept that has received much policy attention in the past decade (see OECD 2001, European Commission 2005). It takes its starting point in the unifying concept of a region which unites various productive forces (capital, labor, infrastructure, knowledge, innovation, resources, or social capital) that act as jointly operating production factors so as to increase regional productivity and efficiency and to sustain socio-economic growth. It may be defined as the set of geographically and locally bounded critical assets, amenities, and conditions that provide the competitive advantages of places and their uniqueness and attractiveness through an efficiency-enhancing contribution to sustainable growth (Camagni 2009). TC is thus not only geographically bounded, but also functionally related to a regional system (in material, social, technological, cultural, and cognitive terms). This idea was inspired by the seminal contributions of Camagni (2002, 2009) , Camagni, Capello (2011, 2013, 2015), Capello et al. (2011), Caragliu (2015), Fratesi, Perucca (2014), and Perucca (2014). This idea has first been introduced as an operational tool in the regional development literature by Camagni (2002) and has been followed by various subsequent empirical territorial capital studies (see for a survey, amongst others, Capello et al. 2011). The three constituents of TC can be briefly described as follows:

- Human capital (HC): knowledge, training, education, R&D, learning-by-doing strategies, creativity, innovativeness, entrepreneurial attitude.
- Infrastructural capital (IC): physical transport and communication facilities, connected networks, energy grids, water facilities, ICT, digital information and moni-

[1] The 'performance' concept already has a long history in industrial management and business economics. In general terms, this concept can be defined as: 'a person's achievement under test conditions' (Oxford Encyclopaedic English Dictionary). However, in productivity and efficiency studies, this concept is defined much more broadly and refers to a systematic operational measurement – often in comparison with relevant actors – of the relevant economic achievement position of an actor or corporate organization (see also Kourtit 2014, 2015, p. 16). The latter meaning will also be adopted in our study on the aviation and aerospace cluster in the Podkarpackie region in South-East Poland.

toring systems.

- Social capital (SC): communication networks, business alliances, incubators, socio-cultural cohesion strategies, cooperation programs, knowledge spillover networks, voluntary organizations (NGOs, etc.).

The smart combination of these productive resources by a region leads to the unifying concept of a 'resourceful region', as advocated in Nijkamp (2016). It is clear that HC, IC, and SC are compound and multidimensional latent vectors, which can only indirectly be observed through measurable indicators (see Section 4). They determine in combination the TPI for each relevant area. In this context, an interesting research question with important regional policy implications concerns the effectiveness of enhancing and improving the regional cluster performance – of both the actors individually and the region as a whole – in producing a desired result on the basis of a smart (innovative and sustainable) use of territorial capital in the area concerned.

In our study we will provide an empirical test of this concept for the Podkarpackie region (voivodship) in South-East Poland. We will address in particular the question: "which territorial capital assets play an efficient role in the mechanism of an effective cluster in the Aviation Valley in this region?" This calls for a triple-layer analysis: (i) a comparative performance analysis of the Podkarpackie region in Poland as the home basis of aviation activities; (ii) a detailed spatial analysis of the counties ('powiats') in Podkarpackie, as the counties ('powiats') in this region are rather heterogeneous and offer different territorial capital assets, and (iii) a comparative study at the firm level to identify the attractiveness factors in the various counties for the individual firms in the Aviation Valley.

3 The Aviation Sector in Poland

3.1 Transformation of the aviation industry in Poland[2]

The airline sector (aviation sector) has over the past decades turned into one of the most dynamic, rapidly growing and high-tech oriented industries in the world, mainly as a consequence of deregulation and intense competition in this sector (see e.g. Adler, Golany 2001, Berechman, de Wit 1996, Burghouwt, Huys 2003, Button 2002, Nijkamp 1996). The demand drivers and the supply conditions of modern aviation have been extensively examined in the literature, in particular, the pricing and routing scheduling, as well as the airport operations. Less attention has been paid to the production and location aspects of aircraft and aircraft equipment and services. The latter – mainly manufacturing – issue has in recent years become an important component of modern regional cluster analysis and policy. Poland has historically been an important center of airplane production, until this sector collapsed after the fall of the iron curtain. Against the background of a dedicated regional policy, the aviation production sector has recently become one of the spearheads of policy support for the Aviation Valley in the Podkarpackie region in South-East Poland.

The history of the Polish aviation industry dates back to the 1930s, when several aviation companies were established in the Central Industrial Area (Centralny Okręg Przemysowy (COP)) in Poland. The transformation of the economic system in Poland in the early 1990's opened new possibilities for the – traditionally strong – Polish aviation industry, but created also several new obstacles that could have slowed down or even stopped its favorable development, if not bypassed effectively. One of the most important strategic decisions was outlining and deciding how the industry should be privatized by choosing and convincing strategic investors (companies which would provide significant investment capital) and by considering and implementing alternative methods of ownership transfer. It was essential to plan how the industry would enter into international cooperative agreements, while showing strength and competitiveness. Market opportunities existing at

[2]The authors wish to thank Zbigniew Bochniarz, Emilia Barbara Sieńko-Kułakowska, Grzegorz Pisarczyk and Waldemar Ratajczak for their great contribution to this section on the transformation and sustainable development of the aviation industry in Poland.

the time indicated that Poland could take advantage of well-established aviation industry facilities and of past achievements, and build up a competitive advantage, particularly in the light-aircraft market. The historical presence of human capital and industrial engineering traditions in the Podkarpackie region – dating back to the pre-WWII period – helped to encourage a revitalization of the aviation industry.

3.2 Sustainable development of the aviation industry in Poland

3.2.1 A sketch of the industry

The expansion of the Polish aviation industry after the transformational period in the 1990s appears to depend strongly on a combination of expansion of international industrial cooperation and foreign direct investment (FDI). As a result, the industry had a rich export offer in terms of advanced aviation products (or rather, sub-assemblies, components, parts, and equipment) for export to e.g., the USA, Venezuela, Indonesia, Italy, Greece, Canada, Spain, Germany, South Korea, and Vietnam. This led to a new and creative revival of industrial spirit and organization. Operating plants in this area are specialized in the production of aircraft components and services (high-tech, training, and executive services), helicopters, gliders, sub-assemblies (aluminum, composite, GRFP), and accessories. Over 140 businesses appeared to operate in 2015 in the Polish aviation and aviation-related sector, with approx. 1 billion EUR in annual sales, and 24,000 employees in total. The majority of these activities are composed of small and medium-sized businesses (SMEs), while companies are partly owned in some cases by foreign investors; there is also a small number of businesses that are owned by the National Treasury in Poland. The majority of the aviation facilities in Poland is located in the South-Eastern part of the country, mainly in a few existing clusters, in which the Aviation Valley in Podkarpackie is the most prominent one.

The biggest foreign investors in the Polish aviation industry are General Electric, United Technologies Corporation, EADS CASA, Pratt & Whitney Canada, Goodrich, Hispano Suiza, and Avio (see Figure 1). The production potential of these businesses is high, mainly due to the quality of the products offered (thanks to a long-standing experience on how to treat materials, casting, mechanics, and electronics) and competitive labor costs. The currently existing network cluster of production and service companies supported by R&D centers have the potential to cooperate in fulfilling and referring orders of replacement parts and complete products for the aviation sector. Thanks to the contribution of companies to R&D, collaboration with research centers and universities, participation in foreign projects, human capital, and the strongly developing clusters, the aviation sector is one of the most innovative ones in the whole Polish economy (see also Ratajczak 2008).

The growth of the aviation industry would not have been possible without qualified human capital. Every year over 11,000 engineers graduate from Polish technical universities, while about 650 graduates have a major in aviation. A highly developed university education system and on-the-job training systems, combined with a rich tradition of excellence, are factors that greatly improve the quality of aviation sector personnel, as is witnessed by the Triple-Helix constellation around the Rzeszow area. Additionally, thanks to new initiatives (such as AERONET), the aviation sector has achieved a close cooperation between industry, government, and educational institutions, serving the common goal of better preparing personnel, for example, by designing specialized educational programs and offering qualified majors to adequately fit the needs of the job market. Clearly, industry-university-policy interfaces appear to be a critical success factor for the growth of the aviation sector in Poland. For a strategy analysis of this phenomenon we refer to Kourtit, Nijkamp (2017).

3.2.2 The Aviation Valley: the Dolina Lotnicza cluster

The Polish aviation industry's prospects for economic and technological advancement are particularly dependent on strategies for effectively operating modern industrial aviation clusters. A promising role model for such a strategy may be found in the Aviation

Source: Internal documents from the Aviation Valley (Dolina Lotnicza) cluster and Marshall Office
Notes: Encircled areas are aviation regions in Poland, in which the Aviation Valley in the South-East is the dominant one

Figure 1: The aviation industry in Poland

Valley (Dolina Lotnicza) cluster in this country. This cluster contains over 100 innovative manufacturing businesses responsible for a large number of aviation products such as components, major assemblies, and sub-assemblies for jet engines, gliders, and helicopters, produced for some of the most important aviation manufacturers and users in the world. Companies in the Aviation Valley are also responsible for the manufacturing of finished products for final clients. Many of the companies in the cluster work together to fulfill their project's tasks in an efficient way, while also a number of businesses complete their entire manufacturing process using only their own resources and offering their products directly to the market. The Aviation Valley represents nowadays about 90% of the aviation industry in Poland. Associated partners are worldwide leaders in their respective fields, such as Pratt & Whitney, Sikorsky, Agusta Westland, Hispano Suiza Polska, Goodrich, MTU Aero Engines, Hamilton Sundstrund, and a dozen of small and medium-sized businesses (see again Figure 1).

A critical challenge in developing this cluster was to bring also small and medium-sized businesses with a commercial link to aviation into the cluster, including those with Polish and foreign investment sources. Through the flow of – and access to – new technologies, these businesses are able to achieve ambitious cooperative goals in the cluster. The realization of such industrial processes means essentially a shift of various activities in the value chain towards the center of the cluster. Experiences from other big clusters indicate that such undertakings lead to effective and high synergy and to measurable economic benefits. The Aviation Valley in Poland has indeed managed to include a large number of critical market and technological leaders among its cluster participants, who continue to bring their added value into the cluster. Consequently, a characteristic of the Aviation Valley cluster is the emergence of a fully developed value chain, as depicted in Figure 2 which shows the structure of the cluster (see also Kaszuba 2012).

The various R&D, production, and market operations in the Aviation Valley cluster have been a model of cooperation between industry and research in Poland. For example, aviation manufacturers meet regularly with representatives of the best Polish technical universities and research centers that have partnered with them. It is noteworthy that among the businesses originating from the Polish aviation industry, only a few large companies (e.g., PZL Mielec, EADS PZL, PZL-Świdnik) produce final goods. Small and medium-sized businesses operate mainly in a cooperative and/or subcontractor role, producing only specific elements of the airframe and equipment of airplanes, or building specialized aircraft parts. However, subcontractors and suppliers are not the only small businesses which play an important role. There are also other small firms that build their own aviation designs, including technologically advanced small aircraft and gliders.

Source: Internal documents from the Aviation Valley (Dolina Lotnicza) Cluster and Marshall Office
Notes: Companies divided by employment level (scale on the left axis)

Figure 2: The structure of the Aviation Valley (Dolina Lotnicza) cluster in Poland

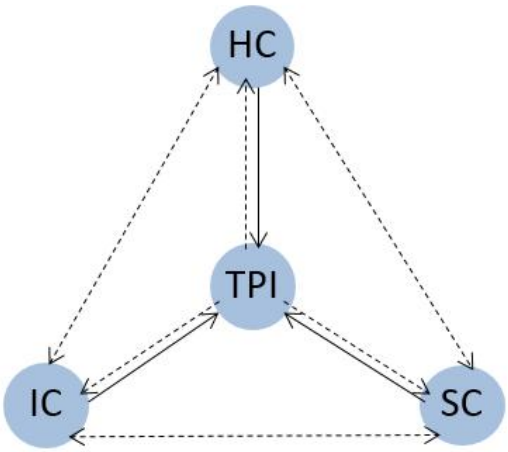
The total sales of all businesses in the aviation sector in Poland have been steadily rising, with the Podkarpackie region as the uncontested leader. The sector has been growing drastically, with sales even quadrupling in the years 2003-2008. This dynamic development is closely linked to foreign investments, as the majority of sales are the result of supplying foreign companies which own or contract these Polish manufacturers.

It is also clear that the presence of and access to an advanced knowledge base in the Podkarpackie area – in particular, higher education institutes – play a critical role in supporting new technology industries (see Calzonetti, Reid 2013). R&D plays a critical role in the success of the Polish aviation industry in general and of Dolina Lotnicza in particular. As a world-class cluster, it has a main goal to conduct advanced academic R&D as well as establishing innovative solutions in the field of aviation engineering (European Aeronautics 2001). Clearly, the Polish aviation industry does currently not yet participate in large-scale manufacturing of passenger planes. This has significant implications for the development of a new model of cooperation with the few world leaders shaping their global supply chain. It influences also the scientific and technological research of the Polish aviation industry and forces it to retain its competitive niche position (Baczko 2011). It is noteworthy that also the Polish job market may be strengthened in the future by a highly qualified workforce able to develop, design, and build high-quality and highly innovative aviation subcomponents and complete products. We will now address the framing of the critical success factors and conditions for this creative aviation development in this region in Poland.

4 A Triple-Layer DEA Model for the Aviation Valley

4.1 Introduction

The notion of an industrial cluster has gained much popularity over the past two decades (see Porter 1990). Recent contributions to cluster concepts and policies can be found inter alia in Asheim et al. (2006), Kasabov, Sundaram (2016), and Scholl, Brenner (2016). As mentioned in Section 1, a cluster may be seen as a complement to earlier concepts from the regional-economic growth literature, such as industrial districts, industrial corridors, growth centers, or development axes (see also a recent contribution by Gibson et al. 2013, Nijkamp 2016). The Aviation Valley – called Dolina Lotnicza in Polish – is a rapidly growing industrial cluster, but needs still further development, in terms of both widening and deepening. In order to become an 'effective cluster' – well anchored in the region on the basis of cooperative strategies ('social capital') and through advanced cognitive and technological applications – in a regional development context, new initiatives are needed, based on entrepreneurship, leadership, and good governance at local and regional levels.



Source: Authors' design

Figure 3: Scheme of the TPI production function (positive externalities for effective cluster performance)

The basic idea of an effective cluster – sketched out in Figure 3 – is that it enhances economic performance of each cluster participant and of the cluster as whole through a smart combination of human, social, and infrastructural capital[3]. Such a cluster is thus driven by positive externalities and may be seen as a flagship for successful regional development.

The Polish aerospace cluster contains more than 100 firms and at present over 1 billion EUR annual revenues, with a concentration in Podkarpackie. This flagship project hosts all major global aviation manufacturers, but still needs a further strengthening and more synergy in order to become an effective and efficient industrial agglomeration, as highlighted in the seminal writings of Marshall, Isard and Porter, to mention only a few.

Industrial clusters are often spontaneously emerging, thematically oriented industrial concentrations, but they may also be the result of dedicated policy decisions to favor a certain industrial concentration in a given region. In both cases, there is a need for informed stimulating policies so as to maximize cluster benefits through innovativeness, productivity increase, and smart specialization. In the specific case of the Aviation Valley, there has been a long tradition of aerospace activities which have been an undercurrent for a rejuvenation of this sector in the past decades after the fall of the iron curtain. Clearly, there is also a need for applied statistical and econometric techniques for testing, understanding and assessing a cluster's performance.

The aim of the present section is to design a systematic and operational framework to provide an original analysis of the cluster achievement data, based on the above mentioned triple-layer architecture, by using an extended and stepwise Data Envelopment Analysis (DEA) to position these regions (voivodships), counties (powiats) and firms on the basis of their relative performance, i.e. by relating their multiple outputs to multiple inputs in the context of an effective cluster to be attained through a balanced mix of HC, IC, and SC, next to the business achievements of individual cluster participants and of the cluster as whole (see also Kourtit, Nijkamp 2013a). We will now offer the design of a conceptual model for an effective cluster assessment in our analysis (see Figure 3).

Figure 3 sketches out the central research aim of the present paper: the assessment of the territorial performance of an advanced industrial area (i.e., an effective cluster) as a function of its human, infrastructural, and social capital provisions (i.e., its total

[3]A comprehensive definition of an effective cluster is: "A geographical and tangible concentration of advanced economic, technological and social activities – both private and public – in a given region, that is driven by synergy coming from network of actors committed (social capital) and usually supported by public policy in order to enhance the cluster performance (in particular, competitiveness, innovation, shared values and trust) – of both the actors individually and the region as a whole – on the basis of a smart (innovative and sustainable) use of territorial capital in the area concerned" (www.effectiveclusters.eu).

territorial capital). The validity and relevance of this TPI methodology will now be tested by means of an application to the Aviation Valley in the Podkarpackie region in South-East Poland, which aspires to become a leading high-tech region in the field of the aviation industry in Europe. We will test our conceptual TPI model on the basis of a triple-layer and a (super-efficient) DEA model using a wide array of empirical data on this Aviation Valley and the territorial capital of the region and of the constituent areas concerned.

It is thus clear that Territorial Capital (TC) is a basket of areal productive assets in an effective cluster composed of Human Capital (HC), Infrastructural Capital (IC), and Social Capital (SC). These three constituents make up the production inputs for the Territorial Performance Index (TPI) in our Data Envelopment Analysis (DEA). Given the large multidimensional databases, the use of a multivariate statistical analysis is a necessary tool in the triple-layer approach in our study so as to obtain a systematically structured database.

4.2 Principal Component Analysis (PCA)

Our triple-layer database contains a wide variety of important statistical factors which determine the quantitative performance of a DMU, at each of the three levels of our analysis (voivodship, powiats and firms). This set of multiple indicators contains at each level an extensive set of multicollinear variables, so that it would be hard to draw straightforward and unambiguous conclusions on the underlying causal mechanisms. Therefore, it was necessary to apply a multivariate analysis – in this case, a principal component analysis – to identify uncorrelated and mutually independent factors, which can be used as proxies for the determinants of regional growth.

In the particular case of detailed statistical information on the 16 voivodships in Poland, we have a broad and detailed annual database (population, labor market, economic variables, etc.) on all voivodships and powiats. Clearly, this leads to a case of multicollinearity in our data. To avoid statistical biases, we used a principal component analysis (PCA), so that we are able to distil from a multicollinear data set a new set of transformed and independent variables that do not suffer from multicollinearity. The next step in our triple-layer analysis was to apply a DEA to each of the three constituents of our database.

4.3 Data Envelopment Analysis (DEA)

DEA has become an established method in management sciences and industrial organization theory to assess the efficiency of complex organizations, often referred to as Decision Making Units (DMUs). It finds its origin in activity analysis and multiple objective programming, and aims to identify in a comparative sense the relative efficiency of DMUs on the basis of their output versus input ratios, in multiple dimensions. So, DEA is essentially a generalized productivity analysis.

The standard DEA model was developed by Charnes et al. (1978) (usually abbreviated as the CCR model). This model – and its many variants and extensions – has found thousands of applications in the scientific literature over the past decades. We refer here to various overview publications, such as Charnes et al. (1994), Zhu (2003), Färe et al. (1998), Ray (2004), Cooper et al. (2006), Zhu, Cook (2007), Suzuki, Nijkamp (2016a,b), Susuki, Nijkamp (2017), and Suzuki et al. (2010, 2015).

The result of a DEA analysis is normally a ranking of DMUs according to their degree of output efficiency. In this way, it is not only possible to find the position of each individual DMU on the efficiency ladder, but also to find out which inputs should be changed, so as to obtain a more efficiently operating organization. There is a great variety of DEA models, starting from the basic CCR model originally developed by Charnes et al. (1978). Over the past decades, a whole range of adjustments and revisions have been implemented, so as to cope with weak elements, limitations, or specific needs of DEA model applications (see Susuki, Nijkamp 2017). We will use here an adjusted version of a standard DEA model, namely a Slack-Based Measure (SBM) – in a triple-layer architecture – to identify the relative efficiency of voivodships, powiats and aviation

business firms in Poland, in order to draw conclusions on the success conditions of the aviation sector in this country and its regions and counties.

In the context of the triple-layer constellation of smart or effective cluster policy in Poland, the so-called Slack-Based Measure (SBM) model turned out to provide a meaningful tool for performing a DEA analysis at three consecutive levels of DMUs, viz. voivodships, powiats, and enterprises. The SBM model was initially developed by Tone (2001) and has found various interesting applications in the literature.

The main distinction between the standard CCR model and the SBM model is related to a radial type projection model and non-radial type model, respectively. A shortcoming of the radial model is a neglect of slacks in computing the efficiency score. Consequently, the radial type model may lead to a biased and overestimated efficiency score. In contrast, the non-radial type models including SBM are able to deal with a slack presence. Hence, an SBM model can improve the overestimation problem.

We will next, in the application of the triple-layer SBM DEA model, include one more extension, viz. the concept of super-efficiency. In the standard DEA model, DMUs located on the efficiency frontier are all equally efficient and hence receive an efficiency score of 1. In many cases, it is desirable to make a further distinction among these efficient firms with an equal score of 1. This has prompted the notion of super-efficiency (SE), through which an unambiguous ranking of DMUs can be achieved (see Tone 2001). We will present here the results of a combined SBM-SE model in our DEA of the Aviation Valley in Poland.

In our empirical DEA application, we will use the (transformed PCA) data on HC, IC, and SC as inputs for a DEA exercise in each of the three stages, while we will use the TPI as output (or performance indicator), based on the architecture of our model sketched in Figure 3. In conclusion, we position PCA and DEA in the context of our triple-layer explanatory model for the performance of relevant DMUs in the aviation industry in the Polish Aviation Valley.

4.4 Conceptual framing of territorial performance in DEA

4.4.1 Architecture of the model

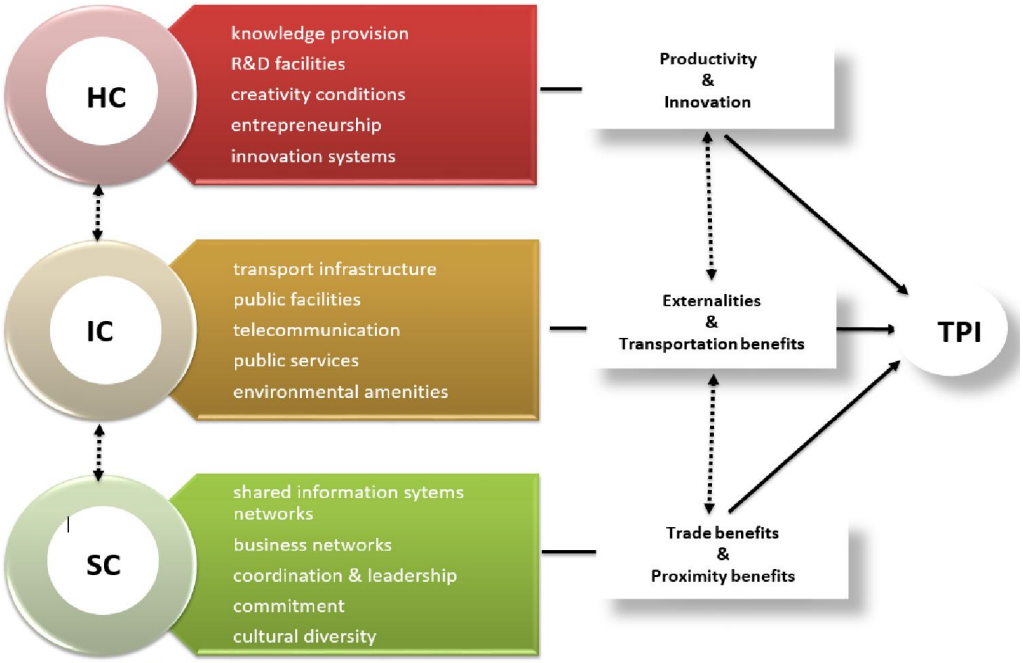
The constituents of the Territorial Performance Index (TPI), viz. HC, IC, and SC, have been outlined in previous sections, based on the schematic presentation in Figure 3. We will, in the present subsection, integrate these building blocks into a more comprehensive TPI framework, by first extending the basic Figure 3 with the most prominent functionalities of the three capital categories under consideration (see Figure 4).

The functional-causal linkages among HC, IC, and SC can be used as the building blocks for a more comprehensive operational model that maps out the total factor productivity (TFP) of the determinants of regional development. A schematic presentation of this model can be found in Figure 4.

It should be added, that next to these contextual capital assets from the business environment in the region, the territorial performance of a region is also determined by the business efforts and subsequent performance actions of private firms in the cluster. The latter category may comprise KPIs of the private sector in the region under consideration. We will now describe in slightly more detail the various explanatory constituents in our TPI production function.

4.4.2 Human Capital

Human capital has become an important topic in economic research since the seminal contributions of Schultz (1961) and later on by Becker (2009). It is nowadays seen as a crucial factor for economic growth and efficiency. Human capital (HC) comprises the personal characteristics and cognitive skills of people who share the responsibility for regional development in a given area. HC is a multi-dimensional concept which includes inter alia: cognitive skills, training facilities, educational programs, advanced knowledge use, R & D expenditure, technical support programs, creativity, human health



Source: Authors' design

Figure 4: A comprehensive representation of the TPI model

conditions, learning experiences, organizing capacity, innovativeness, open innovation systems, entrepreneurship, etc.

4.4.3 Infrastructural Capital

Infrastructure is often regarded as a prominent factor in regional-economic development, in both developed and developing economies (Aschauer 1989, Bröcker, Rietveld 2009, Elburz et al. 2015). Infrastructural capital (IC) refers to the necessary physical and material conditions in the form of public or collective goods that shape or induce the welfare of a country or region. Infrastructure may be instrumental in accelerating economic growth and mitigating spatial disparities. It may adopt various forms, e.g., land transport infrastructure, air and water transport facilities, public amenities and telecommunication. Examples are: roads, streets, parking facilities, railways, railway stations, public transport facilities, airports, runways, ports, energy grids, telecommunication facilities, public services (e.g. health care), environmental amenities, tourist and recreational facilities, etc. The assessment of the impact of IC on regional development calls generally for a broad multidimensional impact model (see e.g., for a general survey and meta-analytical synthesis Celbis et al. 2015).

4.4.4 Social Capital

Social capital is a more recent concept in economic research. It found its origin in sociological research and was first advocated by Hanifan (1916), who described social capital as: "Those tangible assets [that] count for most in the daily lives of people: namely goodwill, fellowship, sympathy, and social intercourse among the individuals and families who make up a social unit" (p. 130). Later on, it was further popularized inter alia by Jacobs (1961), Bourdieu (1981), Coleman (1988) and Fukuyama (2001).

Putnam (1993) offers a more contemporaneous definition: "Social capital refers to features of social organization, such as networks, norms, and trust, that facilitate coordination and cooperation for mutual benefit" (p. 35). Social capital plays a critical role in

Table 1: The multivariate database (reduced into a single-period database with two independent factors by means of a PCA) for the 16 voivodships in Poland[5]

Regions	INPUTS		OUTPUT
	HC	IC	GDP
ŁÓDZKIE	15.24383785	17.76018463	101423
MAZOWIECKIE	59.36306148	1	364513
MAŁOPOLSKIE	27.573995	17.73570387	128009
ŚLĄSKIE	35.9421777	37.32693562	207104
LUBELSKIE	9.357115528	13.42869107	65845
PODKARPACKIE	7.993997384	12.56609439	65365
PODLASKIE	3.351309907	6.687236914	37601
ŚWIĘTOKRZYSKIE	5.554316221	17.32539808	40047
LUBUSKIE	1	11.88567379	36940
WIELKOPOLSKIE	22.30567904	18.32329266	161485
ZACHODNIOPOMORSKIE	9.306254571	13.49000607	62463
DOLNOŚLĄSKIE	22.36283317	14.45307287	140901
OPOLSKIE	4.482978925	20.22856136	35130
KUJAWSKO-POMORSKIE	10.08726379	18.42180737	74515
POMORSKIE	18.02637696	11.99428224	95701
WARMIŃSKO-MAZURSKIE	2.997127294	12.4321011	45008

Source: Source: Authors' own calculation using data from the Statistical Office in Rzeszow (2014)

regional development through the following channels: cooperation among actors, information sharing, trust and honesty, open communication channels, network connectivity, acceptance of coordination or leadership, socio-economic links, cultural bonds and bridges, social commitment, respect for others' values, duty performance, reliability, etc. It goes without saying that the measurement of social capital in an operational explanatory framework is fraught with many difficulties and needs usually the consideration of many variables.

After this description of concepts and data we will now proceed with our analysis and present the database for our triple-layer DEA applications.

5 Information Collection

5.1 Statistical databases

A wealth of statistical information has been collected on regional development in Poland and in the Podkarpackie region, the home base of the Aviation Valley[4]. These data concerned in the first stage the determinants and characteristics of regional welfare in each of the 16 voivodships in Poland in a detailed manner. This macro-regional exploration led to the composition of an extensive multivariate database, with more than 70 indicators comprising five-grained information on the components of Territorial Capital (TC), systematically subdivided into Human Capital (HC), Infrastructural Capital (IC), and Social Capital (SC) (according to the framework in Figures 1 and 2). This multivariate database on HC, IC, and SC indicators was next transformed into a structured database containing two independent factors by means of a principal component analysis (PCA). The endogenous variable employed here (i.e., TPI) is GDP per capita (see Table 1), which is assumed to be determined by the two main components in this table.

The input factors – represented here as two independent vectors after a multivariate

[4]The database for the Aviation Valley study in Poland has been collected in the framework of a research project Effective Clusters – Basis for Innovation and Source of Sustainable Regional Development. The authors wish to thank Krzysztof Kaszuba for his advice on the sources of secondary data, Emilia Barbara Sieńko-Kułakowska and Grzegorz Pisarczyk for their great assistance in collecting primary data, and Waldemar Ratajczak for advice on the various research steps. The detailed source data are contained in the relevant background documents for this project.

Table 2: The multivariate database (reduced into a single-period database with 3 independent factors by means of a FPCA) for the 25 powiats in the Podkarpackie region

POWIATS	INPUTS SC	IC	HC	OUTPUT Average monthly gross wages and salaries (zl)
BIESZCZADZKI	823.34	12.18	84.54	3237.17
BRZOZOWSKI	964.54	59.7	84.66	2964.62
JASIELSKI	713.27	66.86	121.14	3038.82
KROSNIENSKI	954.48	74.88	66.34	2887.05
SANOCKI	584.01	23.32	92.91	3078.45
LESKI	480.87	41.48	121.76	3283.39
M.KROSNO	578.93	281.42	266.59	2980.59
JAROSLAWSKI	535.03	75.96	131.29	3328.63
LUBACZOWSKI	513.1	36.96	93.7	2951.96
PRZEMYSKI	1423.3	43.62	57.41	3036.69
PRZEWORSKI	647.76	87.34	86.93	2931.78
M.PRZEMYSL	487.39	237.84	174.27	3325.5
KOLBUSZOWSKI	939.97	68.46	62.57	2929.58
LANCUCKI	613.6	135.86	97.73	2890.03
ROPCZYCKO-SEDZISZOWSKI	586.34	72.52	103.49	3138.68
RZESZOWSKI	962.87	100.36	74.59	3070.98
STRZYZOWSKI	525.57	62.96	75.23	2786.37
M.RZESZOW	1168.8	343.04	232.39	3859.86
DEBICKI	711.79	89.38	103.39	3165.79
LEZAJSKI	686.66	69.5	104.19	3146.29
MIELECKI	717.79	82.04	114.9	3329.68
NIZANSKI	746.86	61.32	70.86	2890.57
STALOWOWOLSKI	650.9	92.46	122.86	3360.58
TARNOBRZESKI	828.56	81.6	78.74	3326.4
M.TARNOBRZEG	536.3	145.88	153.09	3200.51

Source: Source: Authors' own calculation using data from the Statistical Office in Rzeszow (2014)

transformation from HC, IC, and SC – are the control variables to achieve an output value. Consequently, this data set from stage 1 can be used for an efficiency evaluation by means of DEA.

In the next (meso) layer, viz. the level of the 25 'powiats' (counties) within the Podkarpackie region, a similar approach was adopted, which led again to the composition of three independent factors from SC, IC, and HC, acting as drivers of regional welfare in these counties, while again output or GDP per capita (in terms of average monthly gross wages and salaries) was used as the dependent TPI variable (see Table 2).

Given the rich database at a county level in the Podkarpackie region, in this case the multivariate statistical PCA was separately applied to each of the multidimensional constituents of the HC, IC, and TC indicators in the respective area.

In the third and last (micro) layer, an extensive database was collected on the drivers and perceptions of the individual firms based in the Aviation Valley and located in different powiats in Podkarpackie. This database was the result of semi-structured and often time-consuming detailed interviews with the top managers of these firms, as well as with many other stakeholders. The firms under consideration were both large firms and SMEs. We will now offer a concise description of the elements of the latter database (see also Table 3).

The interview questions covered a wide range of business activities, such as industrial products, forward and backward linkages, service provision, employee types and size, profitability conditions, growth figures, HRM, management structures, links with local

Table 3: Database on drivers and perceptions of individual firms based in the Aviation Valley and located in different powiats in the Podkarpackie voivodship

FIRMS	FA1OU	FA2OU	SC	IC	HC	F1I	F2I
AT1	3	5	650.9	92.46	122.86	2	3
BE2	5	4	1168.8	343.04	232.39	5	4
CA3	4	4	536.3	145.88	153.09	4	3
HE3	2	3	962.87	100.36	74.59	5	2
HI4	5	4	586.34	72.52	103.49	3	4
MA5	5	6	717.79	82.04	114.9	5	3
MT6	5	4	962.87	100.36	74.59	3	4
PZ6	6	6	717.79	82.04	114.9	5	5
TH7	6	4	650.9	92.46	122.86	4	5
TR8	5	6	613.6	135.86	97.73	5	4
UT9	5	3	578.93	281.42	266.59	4	3
WS10	6	6	1168.8	343.04	232.39	5	4
ZM11	5	5	1168.8	343.04	232.39	5	4
AD12	5	5	939.97	68.46	62.57	3	4
AE13	5	6	578.93	281.42	266.59	4	5
AE14	5	6	717.79	82.04	114.9	4	5
AI15	5	5	962.87	100.36	74.59	4	4
AR16	5	0	1168.8	343.04	232.39	3	6
AS17	5	5	1168.8	343.04	232.39	4	4
BM18	4	5	962.87	100.36	74.59	4	5
BO19	5	6	717.79	82.04	114.9	5	4
BR20	6	6	613.6	135.86	97.73	4	5
C021	4	3	1168.8	343.04	232.39	3	4
EL22	5	6	1168.8	343.04	232.39	4	4
EU23	5	5	717.79	82.04	114.9	4	5
FI24	4	1	939.97	68.46	62.57	3	3
FL25	5	5	717.79	82.04	114.9	4	5
FO26	1	5	686.66	69.5	104.19	1	1
GU27	5	6	586.34	72.52	103.49	5	4
HA28	5	5	717.79	82.04	114.9	2	2
IN29	5	6	650.9	92.46	122.86	3	5
IW30	5	6	650.9	92.46	122.86	4	5
KA31	5	0	717.79	82.04	114.9	3	5
MA32	3	5	1168.8	343.04	232.39	3	3
MA33	4	5	717.79	82.04	114.9	3	4
MC34	5	5	962.87	100.36	74.59	4	3
ME35	4	5	962.87	100.36	74.59	3	5
MI36	5	5	578.93	281.42	266.59	2	2
NO37	6	5	1168.8	343.04	232.39	4	4
PO38	4	3	650.9	92.46	122.86	3	3
PO39	5	5	962.87	100.36	74.59	4	5
PZ40	5	5	711.79	89.38	103.39	2	4
RE41	6	6	717.79	82.04	114.9	5	6
R042	5	5	717.79	82.04	114.9	2	4
SE43	6	5	525.57	62.96	75.23	4	4
SP44	4	3	717.79	82.04	114.9	3	4
TW45	5	4	962.87	100.36	74.59	3	1
UN46	4	5	711.79	89.38	103.39	4	4
WA47	5	5	717.79	82.04	114.9	3	5
WI48	6	6	578.93	281.42	266.59	4	4
WS49	4	2	578.93	281.42	266.59	3	3
WY50	5	5	586.34	72.52	103.49	2	3
ZE51	5	5	962.87	100.36	74.59	3	4
ZP52	5	6	717.79	82.04	114.9	5	6

Source: Authors' own calculation using data from interviews in Podkarpackie Voivodship, powiats and gminas (2014) and data of different powiats in the Podkarpackie voivodship from the Statistical Office in Rzeszow (2014)

Notes: Firm names are strictly confidential and therefore anonymized; FA1OU: General growth; FA2OU: Export and revenues; F1I: Strong network and innovation sources; F2I: Cooperation and competitiveness

authorities, cluster linkages, internationalization, etc. The interviews were organized for four distinct groups of business entities of the most important firms and agencies located in the Aviation Valley. We will provide a brief account of these interviews.

The first group consists of large companies which are the main producers for the global airline industry and are closely linked with their parent companies abroad. The second group comprises small and medium-sized companies with a production profile that is partly related to the airline industry and that is linked in a supply chain to large aviation companies in the region. The third group of interviewees addressed institutions of a supporting nature for the aviation business, e.g., universities, consultation and certification institutions, agencies for regional development (all members of the so-called Aviation Valley Association); these organizations are linked to other members of the Aviation Valley Association by providing advisory, training and certificate services, and by jointly carrying out R&D projects. Finally, the last group of interviewees is composed of companies and institutions of a varying legal or administrative status as well as of local and regional authorities related to the Aviation Valley Association, with an external common value chain; examples of such participants are inter alia: institutions with various levels of education – mostly colleges and universities – , banks and financial institutions, local administrations at various levels, other smaller clusters operating in the region, training and consulting companies, and so forth.

As mentioned above, four different types of surveys were designed, while also in-depth survey questions adjusted to specific classes of respondents were added. Clearly, this research task contained a multiplicity of appropriate survey questions geared towards each group of respondents so as to make a comprehensive diagnosis for all respondents in the Aviation Valley whenever possible. In the case of the first two groups of business firms, the survey questions were similar, so that it was possible to compare results which are common to every surveyed company within the cluster, so as to arrive at a consistent interpretation of the results. Our survey questions were designed in a flexible and open way, so that a wide range of specific planning issues for the region could be addressed. Questions were grouped according to a systematic typology of topics, so that the analysis of various questions could provide systematic information on the scope of the research. Clearly, the survey among business firms is to be treated anonymously and its results are used here as summary data without any possibility to identify individual companies which answered particular questions. A total of 55 companies – members of the cluster – took part in the survey.

The primary sources of statistical information were thus supported and complemented by confidential survey research conducted on a group of general companies/institutions cooperating in and with the Aviation Valley Association. The choice of respondents was conditioned by the range of cooperation and the cluster's influence on its economic and technological environment. Therefore, among the institutions and companies interviewed were: (1) financial institutions (banks, insurance companies, support institutions), (2) institutions related to environmental protection and ecology (including informal associations), (3) local government agencies, (4) political parties, (5) energy power companies, (6) transport and logistics companies, (7) local marketing and PR institutions, (8) non-governmental organizations, (9) other local clusters, (10) schools and training institutions, (11) government agencies supporting the cluster activities, and (12) others, e.g., service companies dealing with quality management, human resources management, provision of telecommunication services, and so forth.

The above survey encompassed in total 35 service institutions operating in the Aviation Valley Association environment. It was directed to a total of 73 potential respondents. The response rate for this survey was 48%. The specific results from the group survey questions concerned: general knowledge about the Aviation Valley Association, evaluation of cooperation intensity between the respondent and the cluster (and its members), barriers in the area of business connections, potential motives for entering the cluster, affiliation with other cluster structures and its benefits, evaluation of strategic operations, and marketing and communication of the Aviation Valley Association. Moreover, there were also questions concerning the evaluation of the potential of the region to attract new firms, to further develop a company or the cluster as a whole, or to get access to

sources of development funding or institutional information on who should take action for concrete sustainable development initiatives.

And finally, there were also demographic questions concerning basic information about the human resource profile of the company interviewed.

5.2 In-depth interviews and participatory research

The field research was based on multiple sources of information, from several stakeholders, individual cluster members, local experts, etc. In-depth interviews with cluster participants were also held; they were a crucial element of the field research. An in-depth interview is essentially a communication exercise with a respondent conducted on the basis of a pre-specified mental scenario. Interviews were conducted by persons especially trained for this aim as part of the project, while the interviewer was an active and qualified participant in the communication, in order to gather meaningful strategic and operational information. In-depth interviews are an alternative to focus-group interviews. Due to the specific aim of the project and the diversity of respondents, direct in-depth interviews were chosen. They provide in general a better feeling of ease among interviewees. This is especially important when it comes to issues related to trust evaluation and cooperation between representatives of other companies and institutions within the cluster.

For the sake of comprehensiveness of the research issues, respondents in the cluster were chosen according to following criteria: hierarchal position in the company or institution surveyed, membership of the Aviation Valley Association, and involvement in operations initiated by the Aviation Valley Association. The interviews were carefully prepared in advance, while relevant information about the aim of the project and of the research and about the nature of the questions to be raised during the interview was provided. For representatives of big companies, small and middle-sized companies, and institutions of the business environment who were members of the cluster, similar questions were prepared.

Finally, the aim of the participatory research was to engage representatives of the cluster members in the decision processes and to help in shaping the cluster's development policy. The aim of this action research was to gain knowledge about the future business environment, and strategic knowledge oriented towards the solution of pressing problems. Both aims are interrelated and therefore, it was necessary to engage participants in the organization, as well as in the formulation of strategic conclusions regarding the diagnosis of their collective interests and their connections. The aim of engaging key actors of the cluster in participatory research was also to create awareness of essential issues for these actors, and to show that their voice counts, so that they have a real impact on what is going to happen in the cluster's future. Such research may become a catalyst of proactive involvement, and may create the foundation of establishing participatory mechanisms in the cluster.

6 Results of a Triple-Layer DEA Benchmark Analysis in the Podkarpackie Region in Poland

The predominant aim of our research is now to assess the relative economic performance of Polish regions, of counties inside the main region concerned (Podkarpackie), and of individual firms in the Aviation Valley on the basis of an efficiency (or productivity) analysis by means of a DEA (see Figures 3 and 4). In our empirical application we will use the above presented extensive database on Polish regions, counties, and individual actors in the aviation cluster, in combination with the rich performance database for the Podkarpackie region in Poland. We are seeking to achieve a ranking of agents in the triple-layer architecture based on a comprehensive set of indicators and aim to assess the efficiency (or productivity) of the different layers of agents in the context of an effective cluster performance, by examining more carefully the ratio between multi-attribute outputs and multi-attribute inputs of each layer. As mentioned, data envelopment analysis (DEA) is an appropriate tool in this context. Thus, our study aims to provide a critical analysis of the performance data of DMUs in the triple-layer architecture of the Aviation Valley by using an SBM model and a super-efficient (SE) DEA approach, respectively, to position

Table 4: SBM-SE scores of 16 NUTS-3 Voivodships of Poland

Nr.	DMU Voivodship	Score	Nr.	DMU Voivodship	Score
1	DOLNOŚLĄSKIE	0.561	9	PODKARPACKIE	0.755
2	KUJAWSKO-POMORSKIE	0.636	10	PODLASKIE	0.995
3	ŁÓDZKIE	0.592	11	POMORSKIE	0.443
4	LUBELSKIE	0.626	12	ŚLĄSKIE	0.477
5	LUBUSKIE	1.730	13	ŚWIĘTOKRZYSKIE	0.469
6	MAŁOPOLSKIE	0.388	14	WARMIŃSKO-MAZURSKIE	0.849
7	MAZOWIECKIE	19.195	15	WIELKOPOLSKIE	0.768
8	OPOLSKIE	0.385	16	ZACHODNIOPOMORSKIE	0.582

Source: Authors' own calculation

Table 5: SBM-SE scores of 25 NUTS-4 Powiats of Podkarpackie in Poland

Nr.	DMU Powiat	Score	Nr.	DMU Powiat	Score
1	bieszczadzki	1.390	13	m.Przemysl	1.003
2	brzozowski	0.687	14	m.Rzeszow	0.642
3	debicki	0.908	15	m.Tarnobrzeg	0.727
4	jaroslawski	1.012	16	mielecki	0.708
5	jasielski	0.693	17	nizanski	1.082
6	kolbuszowski	1.062	18	przemyski	0.628
7	kroSnieNski	0.711	19	przeworski	0.757
8	lancucki	0.679	20	ropczycko-sedziszowski	0.628
9	leski	0.637	21	rzeszowski	0.649
10	lezajski	0.404	22	sanocki	1.038
11	lubaczowski	0.788	23	stalowowolski	1.039
12	m.Krosno	1.011	24	strzyzowski	0.811
			25	tarnobrzeski	0.434

Source: Authors' own calculation

these DMUs unambiguously on the basis of their relative performance. The empirical results of the SBM-SE model based on the triple-layer approach will now be presented in Tables 4-6 for each of the three layers concerned.

We will offer here a concise interpretation of the findings in these tables. As far as Table 1 is concerned, it turns out that there are only 2 super-efficient voivodships on the NUTS-3 level in Poland, with a clearly prominent position for Mazowieckie, situated in the central-eastern part of Poland. There is quite some variation among the efficiency outcomes of the non-efficient voivodships in Poland. It is noteworthy that the efficiency of the Podkarpackie region – which is the region addressed in our study – does not differ significantly from the Wielkopolskie region which is a rather strong region in the Western part of Poland dominated by Poznan. Thus, the Podkarpackie region is performing rather well, despite its peripheral location. This may be seen as a promising sign for future development efforts.

On the next layer of powiats (NUTS-4 regions or counties) within the Podkarpackie region, we observe quite a few (8) super-efficient areas. They are scattered all over the Podkarpackie voivodship, which means that the efficiency performance of these counties offers a balanced geographical picture, except for these powiats that are external border regions (e.g., to Ukraine). This finding is interesting, as the aviation industry in the Podkarpackie region is not concentrated in one point location, but shows a broadly dispersed cluster pattern all over this region. Apparently, the benefits of this strategy accrue to a wide group of counties in the voivodship. The relatively low score of Rzeszow has to do with the fact that the city itself does not have a strong industrial base, since most aviation activities are located in a wide radius around the city.

Table 6: SBM-SE scores of 55 firms in the Aviation Cluster in Podkarpackie in Poland

Nr.	DMU Firms	Score	Nr.	DMU Firms	Score
1	AD10	1.016	28	MA39	0.710
2	Ae11	1.007	29	MC27	0.731
3	Ae35	0.829	30	Me14	1.000
4	Ai20	0.844	31	MI52	0.316
5	AR4	1.000	32	MT30	0.744
6	AS48	0.502	33	No50	0.475
7	AT40	0.559	34	Po15	0.960
8	BE42	0.601	35	Pol43	0.691
9	BM8	1.049	36	PZ19	0.927
10	Bo21	0.916	37	PZ45	0.646
11	BR17	0.793	38	Re16	0.967
12	CA13	1.001	39	Ro47	0.640
13	CO38	0.476	40	SE7	1.032
14	EL48	0.493	41	Sp28	0.775
15	Eu32	0.857	42	Th18	0.862
16	Fi3	1.053	43	Tr12	1.008
17	FL33	0.857	44	TW51	0.498
18	Fo36	0.347	45	Un29	0.830
19	GU5	1.088	46	UT9	1.022
20	HA53	0.396	47	Wa31	0.800
21	He2	1.529	48	Wi26	0.582
22	Hi34	0.767	49	WS25	0.613
23	In24	0.805	50	WS46	0.540
24	IW23	0.816	51	Wy49	0.545
25	KA1	1.000	52	Ze41	0.714
26	Ma22	0.887	53	ZM44	0.581
27	Ma37	0.456	54	ZP6	1.040

Source: Authors' own calculation

Finally, the efficiency scores for the firms in the aviation industry shows an interesting pattern. There is apparently a set of 16 super-efficient firms, and a broad distribution of less efficient firms. As mentioned, details on these anonymized firms cannot be provided, but there is also a fair balance between small and large firms. It is also interesting to observe that there are a limited number of very inefficient firms with rather low scores. Consequently, for specific enterprises in this aviation cluster there is much scope for improving their business performance. It is also noteworthy that the geographic distribution of efficient and inefficient firms over the various powiats is rather balanced.

The overall conclusion from our triple-layer DEA model application is that on all levels of decision-making – ranging from individual business firms in the aviation cluster through counties (which provide the direct geographical location area for these firms) to the level of voivodships charged with official regional policy competences – there is still much scope for socio-economic performance improvement. The Podkarpackie region has in the past decade demonstrated a great potential in enhancing its competitive position as an effective aviation cluster (see Kaszuba 2012), but there is undoubtedly more room for strengthening its position.

It should be added that in a deterministic DEA model these findings are of course numerically correct, but that in policy practice such accurate and ambitious adjustments by DMUs may sometimes be hard to achieve. Nevertheless, this information is indicative for the direction and intensity of necessary policy handled in a region, powiat, and firm so as to become more efficient in a competitive environment. In conclusion, there is a need for an intensified policy effort to achieve and strengthen the socio-economic vitality and efficiency of the aviation actors in the triple-layer architecture in Poland.

7 Policy Lessons

Regional development policy aims to cope with the challenge of spatial disparities. It is based on a smart combination of various critical capital assets in a region which functionally and spatially interact and which yield synergetic economic opportunities and promising challenges for innovation and progress. The present study regards sustainable territorial performance – as a manifestation of regional development – as the overarching principle for competitive advantages and economic growth in a system of regions, which is particularly induced by territorial capital, comprising human capital, infrastructural capital, and social capital. In the long-standing tradition of regional development policy, a wide variety of effective facilitators or drivers of accelerated spatial growth has been distinguished, for instance, industrial districts, growth poles, growth centers, industrial complexes, special economic zones, communication axes, and so forth.

Regional development calls for an active involvement of stakeholders or agents (DMUs) at different institutional levels. Awareness of the actual position of a DMU on the achievement ladder is a sine qua non for performance improvement. The Aviation Valley cluster in the Podkarpackie region in South-East Poland offers a good example of a dedicated but complex regional development effort; it is not only a fascinating and pioneering cluster policy experiment, but offers also a great opportunity for designing and implementing innovative and effective cluster research initiatives and for building strong business networks and creating 'collective buying power'. It is indeed a miraculous and encouraging phenomenon that a less privileged and peripheral territory like the Podkarpackie region is able to act as a seedbed for a broad portfolio of advanced aviation activities, not only with a great economic spinoff for the area concerned, but also with significant spillovers to the rest of the country and of Europe, and in various cases even with a world-wide outreach.

The Podkarpackie voivodship and its powiats are confronted with a great many challenges, notably geographic isolation, far-reaching demographic transformations, complex force fields in this part of Europe, unequal social participation, and ever-rising mobility trends. These multi-faceted challenges may be turned into new opportunities, in particular in such concerns as advanced business and environmental facilities, knowledge-intensive and creative strategies for socioeconomic well-being and prosperity (see Kourtit, Nijkamp 2013b) with a strong need for intensified policy efforts, and availability of a strong and dedicated workforce in high-tech engineering.

The general challenge is to improve the competitive high performance of the Podkarpackie region, to strengthen and stimulate its constituent powiats, and to improve further the attractiveness of the Aviation Valley as a sustainable high-quality place to work and live, and to incorporate technology and innovation in overall sustainable developmental strategies so as to make the Valley a seedbed for global frontrunners for future development. The notion of a territorial performance index (TPI) has proven its relevance and applicability in our research.

In this context, strong and fit-for-purpose territorial capital may create a strong Aviation Valley, in particular, through strengthening information and knowledge facilities, advancing institutional support systems, fostering fruitful business networking initiatives, recruiting new talents on the labor market, favoring high-skilled job opportunities and establishing a sufficient and effective venture capital system (see Kourtit, Nijkamp 2013b, Kourtit 2015). All such effective cluster conditions are critical success factors for a promising living, working, and business environment in the Aviation Valley. A mixture of advanced process, product and service innovation initiatives, advanced labor force concentrations, socio-cultural initiatives, interconnected public facilities, and geographic knowledge synergy may be regarded as the constituents of a successful regional creative and innovation system that forms the basis for an effective aviation cluster.

In our applied empirical case study on the Aviation Valley in the Podkarpackie region in Poland, we have focused our attention in particular on the quality of human capital, social characteristics, and on infrastructural amenities synergy from the perspective of an effective cluster against the background of social and economic dynamics at different scale levels, on the basis of the TPI model. Clearly, the aim of our analysis was not to

offer a critical review of current official development policies and implementations in the region concerned. We have rather addressed the information base and the conceptual framing of the strong and weak points in this area, with a particular view to regional, local, and industrial detail.

At the level of Polish voivodships, the regions Mazowieckie and Lubuskie are found to be efficient (with efficiency scores equal to 1) based on our DEA model. This "efficiency" means that these two regions can produce a large amount of economic performance (outputs), with the use of relatively small amount of inputs (territorial capital). This result may be confronted with relatively highly inefficient regions, such as Małopolskie, Świętokrzyskie and Opolskie. The Podkarpackie region is found in a middle position with a trend to improve its performance.

At the next spatial scale, viz. of powiats, the most efficient powiats in the Podkarpackie region in Poland are Bieszczadzki, Sanocki, Leski, Lubaczowski, Przemyski, Kolbuszowski, Strzyzowski and Tarnobrzeski. These powiats appear to offer new opportunities for social and economic synergy as a result of a strong county attractiveness and the presence of broadly based public and private facilities. There is obviously also a great opportunity for advanced growth initiatives and improvement strategies of lower performing powiats. Apparently, territorial capital is very supportive for high aviation industry performance.

Finally, by looking at the third stage for efficient firms in the Aviation Valley in the Podkarpackie region in Poland, we can by means of our DEA model identify the firms with a maximum level of efficiency, which from our analysis turns out to comprise 16 firms. These firms may be labelled the 'high performing firms' among our sample of 55 high-tech firms in the Aviation Valley. These 'high performing firms' are able to exploit their high socio-economic and technological performance even more, as they have strongly established business resources that underscore the importance of territorial capital.

In conclusion, a strategic view on a high performing region, country, or firm may be used as an effective approach to identify important and smart KPIs (key performance indicators) that are involved to maximize the creative, innovative, and technological potential of the agent concerned. The triple-layer design of our research, addressing three types of actors with distinct competences, has proven to be a valuable methodological departure for an effective cluster analysis of the Aviation Valley in Poland. This approach may provide untapped opportunities for knowledge-based institutional, geographically-focused, and cultural and business activities, that are needed in order to stay ahead of the fierce competition and to advance continuously the high socio-economic achievement potential.

Clearly, any DEA study – including ours – is a comparative benchmarking analysis of the efficiency – or the broad economic performance – of various stakeholders or actors (e.g., regions, agencies, firms). The reasons why some actors have an exceptional performance in comparison to others in the same sample call for further specific case research. The same holds for all other DMUs who want to improve their performance position on the achievement ladder. It should be added that a sensitivity analysis using another DEA method, viz. the CCR model, leads to largely the same result, so that our findings seem to be rather robust and consistent. It should be added that essentially a wide array of DEA variants could have been applied (such as BCC, Malmquist indices, etc.), but that would have led to a different study. For more details on this option we refer to a recent study by Susuki, Nijkamp (2017).

The overall conclusion is that the Aviation Valley in the Podkarpackie region in Poland has unique opportunities and a rich scope for further improvement and development and may act as a catalyst for welfare improvement in this territory.

References

Acz Z, Braunerhjelm P, Audretsch D, Carlsson B (2009) The knowledge spill-over theory of entrepreneurship. *Small Business Economics* 32: 15–30. CrossRef.

Adler N, Golany B (2001) Evaluation of deregulated airline networks using data envelopment analysis combined with principal component analysis with an application to Western Europe. *European Journal of Operational Research* 132[2]: 260–273. CrossRef.

Andersson M, Klaesson J, Larsson JP (2016) How local are spatial density externalities? Neighbourhood effects in agglomeration economies. *Regional Studies* 50[6]: 1082–1095. CrossRef.

Arribas-Bel D, Kourtit K, Nijkamp P (2013) Benchmarking of world cities through self-organizing maps. *Cities* 31: 248–257. CrossRef.

Aschauer DA (1989) Is public expenditure productive? *Journal of Monetary Economics* 23: 177–200. CrossRef.

Asheim B, Cooke P, Martin R (2006) *Clusters and Regional Development: Critical Reflections and Explorations*. Routledge, London

Baczko T (2011) Report on innovativeness of the aviation sector in Poland in 2010. http://issuu.com/inepan/docs/ang_pdf_raport_okladki

Batabyal A, Nijkamp P (2015) *The Region and Trade: New Analytical Directions*. World Scientific Publishing, New York

Becker GS (2009) *Human Capital*. University of Chicago Press, Chicago. CrossRef.

Berechman J, de Wit JG (1996) An analysis of the effects of European aviation deregulation on an airline's network structure and choice of a primary West-European hub airport. *Journal of Transport Economics and Policy* 30[3]: 251–274

Bochniarz Z (2007) Strategic issues of financing sustainable development in transforming economies: The case of Central Europe. In: Xiaowen J, Ebben J (eds), *Entrepreneurial Strategy Innovation and Sustainable Development*. Sichuan University Press, Chengdu, China, 1029–1049

Bourdieu P (1981) Décrire et préscrire. note sur les conditions de possibilité et limites de l'efficacité. *Actes de la Recherche en Sciences Sociales* 38[1]: 69–73. CrossRef.

Bröcker J, Rietveld P (2009) Infrastructure and regional development. In: Capello R, Nijkamp P (eds), *Handbook of Regional Growth and Development Theories*. Edward Elgar, Cheltenham, 152–281. CrossRef.

Burghouwt G, Huys M (2003) Deregulation and the consequences for airport planning in Europe. *disP-The Planning Review* 39[154]: 37–45. CrossRef.

Button K (2002) Debunking some common myths about airport hubs. *Journal of Air Transport Management* 8[3]: 177–188. CrossRef.

CAI – Community Attributes Inc. (2013) Washington State Aerospace Industry. Community Attributes Inc. (CAI), Seattle, Washington

Calzonetti FJ, Reid N (2013) The role diversities in supporting new technology industries through commercialisation spin-off activities. *Studies in Regional Science* 43[1]: 7–23. CrossRef.

Camagni R (2002) On the concept of territorial competitiveness: sound or misleading? paper presented at the ERSA conference 2002, European Regional Science Association, http://www-sre.wu.ac.at/ersa/ersaconfs/ersa02/cd-rom/papers/518.pdf

Camagni R (2009) Territorial capital and regional development. In: Capello R, Nijkamp P (eds), *Handbook of Regional Growth and Development Theories*. Edward Elgar, Cheltenham, 118–132. CrossRef.

Camagni R, Capello R (2011) *Spatial Scenarios in a Global Perspective: Europe and the Latin Arc Countries*. Edward Elgar, Cheltenham. CrossRef.

Camagni R, Capello R (2013) Regional competitiveness and territorial capital: A conceptual approach and empirical evidence from the European Union. *Regional Studies* 47[9]: 1383–1402. CrossRef.

Camagni R, Capello R (2015) Rationale and design of EU cohesion policies in a period of crisis. *Regional Science Policy and Practice* 7[1]: 25–49. CrossRef.

Capello R, Caragliu A, Nijkamp P (2011) Territorial capital and regional growth: Increasing returns in knowledge use. *Tijdschrift voor Economische en Sociale Geographie* 102[4]: 385–405. CrossRef.

Capello R, Nijkamp P (2009) *Handbook of Regional Growth and Development Theories*. Edward Elgar, Cheltenham. CrossRef.

Caragliu A (2015) The economics of proximity. PhD dissertation, VU University, Amsterdam

Celbis G, Nijkamp P, Poot J (2015) Infrastructure and trade: A meta-analysis. *REGION* 1[1]: 25–64. CrossRef.

Charnes A, Cooper WW, Lewin AY, Seiford LM (1994) *Data Envelopment Analysis: Theory, Methodology and Applications*. Kluwer, Boston. CrossRef.

Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research* 2: 429–444. CrossRef.

Coleman JS (1988) Social capital in the creation of human capital. *The American Journal of Sociology* 94: 95–120. CrossRef.

Cooper WW, Seiford L, Zhu J (2006) *Handbook of Data Envelopment Analysis*. Kluwer, Boston. CrossRef.

De la Fuente A, Vives X (1995) Infrastructure and education as instruments of regional policy. *Economic Policy* 20: 11–54

ECORYS (2009) Competitiveness of the EU aerospace industry. Rotterdam

Elburz Z, Nijkamp P, Pels E (2015) Public meta-analysis. Discussion paper, Department of Spatial Economics, VU University, Amsterdam

European Aeronautics (2001) Meeting society's needs and winning global leadership. Report of the group of personalities, http://www.acare4europe.org/sites/acare4europe.org/files/document/Vision%202020_0.pdfm

European Commission (2005) Territorial state and perspectives of the European Union. Scoping document and summary of political messages, Brussels, Belgium

Fratesi U, Perucca G (2014) Territorial capital and the effectiveness of cohesion policies: An assessment for CEE regions. *Investigaciones Regionales* 29: 164–191

Färe R, Grosskopf S, Russell R (1998) *Index Numbers*. Kluwer, Boston

Fukuyama F (2001) Social capital, civil society and development. *Third World Quarterly* 22[1]: 7–20. CrossRef.

Gibson MD, Kundu S, Satish M (2013) Dispersion model evaluation of PM2.5, NOx and SO2 from point and major line sources in Nova Scotia, Canada using AERMOD Gaussian plume air dispersion model. *Atmospheric Pollution Research* 4: 157–167. CrossRef.

Hanifan LJC (1916) The rural school community centre. *Annals of the American Academy of Political and Social* 67[1]: 130–138. CrossRef.

Healy P, Cote S (2001) The wellbeing of nations: The role of human and social capital, education and skills. Oecd, paris

Heckman JJ (2000) Policies to foster human capital. 54[1]: 3–56. CrossRef.

Jacobs J (1961) *The Death and Life of Great American Cities*. Random House, New York. CrossRef.

Kaasa A, Part A (2008) Human capital and social capital as interacting factors of economic development: Evidence from Europe. IAREG, Estonia, working paper IAREG WP 2/04. CrossRef.

Kasabov E, Sundaram U (2016) Conceptualizing clusters as dynamic and path-dependent pools of skills. *Regional Studies* 50[9]: 1520–1536. CrossRef.

Kaszuba K (2012) Aviation cluster 'aviation valley' - case study. Report, Rzeszow School of Business, Rzeszow, Poland

Kourtit K (2014) Competiveness in urban systems - studies on the 'urban century'. VU University, Amsterdam, The Netherlands

Kourtit K (2015) The 'new urban world' - economic-geographical studies on the performance of urban systems. Adam Mickiewicz University, Poznan, Poland

Kourtit K, Nijkamp P (2013a) In search of creative champions in high-tech spaces: A spatial application of strategic performance management. *Journal of Regional Science* 53[5]: 749–777. CrossRef.

Kourtit K, Nijkamp P (2013b) The new urban world - The challenge of cities in decline. *Romanian Journal of Regional Science* 7[2]: 9–28

Kourtit K, Nijkamp P (2017) A multi-layer analysis of effective clusters in regional development policy: A case study on Poland. Public administration and regional studies (forthcoming)

Kourtit K, Nijkamp. P, Suzuki S (2013) Exceptional places: The rat race between world cities. *Computers, Environment and Urban Systems* 38: 67–77

Nijkamp P (1996) Liberalization of air transport in Europe: The survival of the fittest? *Swiss Journal of Economics and Statistics* 132[3]: 257–278

Nijkamp P (2008) XXQ factors for sustainable urban development: A systems economics view. *Romanian Journal of Regional Science* 2[1]: 1–34

Nijkamp P (2016) The resourceful regions: A new conceptualisation of regional development strategies. *Journal of Regional Research* 36: 191–214

OECD (2001) OECD territorial outlook. OECD, Paris

Perucca G (2014) The role of territorial capital in local economic growth: Evidence from Italy. *European Planning Studies* 22[3]: 537–562. CrossRef.

Porter ME (1990) *The Competitive Advantage of Nations*. MacMillan, London. CrossRef.

Putnam R (1993) The prosperous community. *The American Prospect* 13[4]: 35–42

Ratajczak W (2008) Impact of Poland's integration with the European Union on the socio-economic status of Polish regions. *Moravian Geographical Reports* 16[1]: 2–15

Ravikumar B, Glomm G (1992) Public versus private investment in human capital. *Journal of Political Economy* 100[4]: 818–834. CrossRef.

Ray SC (2004) *Data Envelopment Analysis: Theory and Techniques for Economics and Operations Research*. Cambridge University Press, Cambridge

Rodriguez-Pose A, Fratesi U (2004) Between development and social policies. *Regional Studies* 38: 97–113

Schönfeld T, Jouaillec F (2008) The regional aerospace cluster policy in Europe. Paper presented to the 26th International Congress of the Aeronautical Sciences. http://-www.icas.org/ICAS_ARCHIVE/ICAS2008/PAPERS/628.PDF

Scholl T, Brenner T (2016) Detecting spatial clustering using a firm-level cluster index. *Regional Studies* 50[6]: 1054–1068. CrossRef.

Schultz Th (1961) Investment in human capital. *American Economic Review* 51[1]: 1–17

Sommers P, Beyers WB, Wenzel A (2008) Industry cluster analysis for Washington State, workforce development areas. Washington State Workforce Board, Seattle, Washington

Statistical Office in Rzeszow (2014) Podkarpackie voivodship – subregions, powiats, gminas, podkarpackie voivodship. http://rzeszow.stat.gov.pl/publikacje-i-foldery/roczniki-statystyczne/wojewodztwo-podkarpackie-podregiony-powiaty-gminy-2014,4,11.html

Susuki S, Nijkamp P (2017) An evaluation of energy-environment-economic efficiency for Asian countries: A proposal for a time-series target-oriented DFM model in data envelopment analysis. In: Batabyal A, Nijkamp P (eds), *Regional Growth and Sustainable Development in Asia*. Springer, Singapore (forthcoming). CrossRef.

Suzuki S, Nijkamp P (2016a) An evaluation of energy-environment-economic efficiency for EU, APEC and ASEAN countries: Design of a target-oriented DFM model with fixed factors in data envelopment analysis. *Energy Policy* 88: 100–112. CrossRef.

Suzuki S, Nijkamp P (2016b) Preference elicitation in generalized data envelopment analysis: In search of a new energy balance in Japan. In: Shibusawa H, Sakurai K, Mizunoya T, Uchida S (eds), *Socioeconomic Environmental Policies and Evaluations in Regional Science*. Springer, Singapore, 601–618. CrossRef.

Suzuki S, Nijkamp P, Rietveld P (2015) A target-oriented data envelopment analysis for energy-environment efficiency improvement in Japan. *Energy Efficiency* 8[3]: 433–446. CrossRef.

Suzuki S, Nijkamp P, Rietveld P, Pels E (2010) A distance friction minimization approach in data envelopment analysis: A comparative study on airport efficiency. *European Journal of Operational Research* 207: 1104–1115. CrossRef.

Tone K (2001) A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research* 130: 498–509. CrossRef.

Westlund H (2006) *Social Capital in the Knowledge Economy: Theory and Empirics*. Springer, Berlin, Heidelberg

Zhu J (2003) *Quantitative Models for Performance Evaluation and Benchmarking*. Springer, Berlin. CrossRef.

Zhu J, Cook W (2007) *Modeling data irregularities and structural complexities in data envelopment analysis*. Springer, New York

Zuliani JM, Jalebert G (2005) L' Industrie Aeronautique Européenne. *L' Espace Geographique* 34[2]: 117–144

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 3, 2017, 65–76
DOI: 10.18335/region.v4i3.167

journal homepage: region.ersa.org

Panel Data Models of New Firm Formation in New England

Jitendra Parajuli[1], Kingsley E. Haynes[2]

[1] George Mason University, Arlington, VA, USA
[2] George Mason University, Arlington, VA, USA and Hong Kong University of Science and Technology, Kowloon, HK, China

Received: 10 October 2016/Accepted: 9 August 2017

Abstract. This study examines the impact of the determinants of new firm formation in New England at the county level from 1999 to 2009. Based on the Spatial Durbin panel model that accounts for spillover effects, it is found that population density and human capital positively affect single-unit firm births within a county and its neighbors. Population growth rate also exerts a significant positive impact on new firm formation, but most of the effect is from spatial spillovers. On the contrary, the ratio of large to small firms in terms of employment size and the unemployment rate negatively influences single-unit firm births both within counties and among neighbors. However, there is no significant impact of local financial capital and personal income growth on new firm formation.

1 Introduction

Entrepreneurship is an important component of the US economy. Schramm (2004) suggested that the US is the leading entrepreneurial country in the world. New firms drive innovation and even during recessionary periods entrepreneurs provide impetus for recovery and economic growth. More recently, the US was ranked as the most entrepreneurial nation among 132 countries in the world and it was suggested that the US has entered a new innovation-driven stage of development (Acs et al. 2015).

Although the National Bureau of Economic Research (NBER) noted that the US experienced major economic contractions in 2001 and from late 2007 through mid-2009 (NBER 2009), the rate of new entrepreneurs (i.e., the percentage of adult population becoming entrepreneurs) has not changed abruptly. In 2001, this rate was 0.28% and from 2007 through 2009, it changed from 0.30% to 0.34% (Fairlie et al. 2016). However, such an aggregate index does not reflect the actual variation of entrepreneurial activities at the regional and local level. Moreover, spatial dependence and heterogeneity often mask local spatial patterns and spillovers (see Anselin 1988, LeSage, Pace 2009). This paper argues that while new firm formation is determined by the local characteristics of individual regions over time, it is also important to understand the effect of spatial externalities in firm births. Using fixed effects panel data models, this study examines the impact of the determinants of new firm formation in the high innovation region of New England – Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont. This is done at the county level from 1999 through 2009 in order to capture local effects.

This paper is arranged in the following sections. A theoretical background is presented in the next section (Section 2). Research methodology, data description, and information sources are presented in Section 3. The empirical results are provided in Section 4. Section 5 includes concluding remarks, policy recommendations, and directions for future research.

2 Literature Review

For various reasons, economic activities are not consistent across space and time. Following Marshall, Krugman (1991) argued that businesses agglomerate in a region because of the pooled labor market, production of non-tradable specialized inputs, and the possibility of greater productivity due to information spillovers. Agglomeration (Ellison, Glaeser 1999), area-based policy (Anyadike-Danes, O'Reilly 2005), cluster strategy (Delgado et al. 2010), and regional embeddedness (Dahl, Sorenson 2012) all influence business location decisions and the distribution of economic activities.

Population density is an important determinant of new firm births. According to Reynolds et al. (1994), urbanization and agglomeration are closely associated with population density and new firm formation rates are often positively associated with population density. Audretsch, Fritsch (1994) also found a positive relationship between population density (agglomeration) and new firm births. The percentage of entrepreneurial activity is higher in urban regions that are characterized by high population density (Bosma, Schutjens 2011). However, if a region has already maximized the benefits of urbanization, high population density can have negative impacts as well (Delfmann et al. 2014).

Human capital and entrepreneurship are closely associated with each other (Garvin 1983, Robinson, Sexton 1994, Davidsson, Honig 2003). Fritsch (1992) suggested that new business founders are highly skilled and Armington, Acs (2002) noted the importance of human capital on firm births. Regional variation in new firm formation, especially in the service sector, is contingent upon the availability of college-educated individuals that normally establish and manage new business ventures (Acs, Armington 2004). Likewise, human capital is crucial for innovation and information flow, and hence promotes new firm formation (Lee et al. 2004).

Financial capital is crucial for starting a new business. Personal, informal – acquaintances and angel investors – and/or formal financial institutions, such as banks, are often the sources of investment. Often, new businesses and entrepreneurs reap the benefits of the local financial environment where they were born or have lived for long time periods (Michelacci, Silva 2007, Robinson, Cottrell 2007, Dahl, Sorenson 2012). Sutaria, Hicks (2004) found a positive relationship between local bank deposits per capita and new firm formation. On the contrary, Kim et al. (2006) noted that potential entrepreneurs gain more from human capital and that financial capital is not a necessary condition for entrepreneurial ventures.

Establishment size is equally important for new firm formation. Audretsch, Fritsch (1994) examined the relationship between new firm births and mean establishment size using different frameworks and found both positive and negative association between them. In general, there is a negative association between average firm size and births in a region. That is, the larger the mean size of existing firms, the smaller the rate of new firm formation (Kangasharju 2000, Armington, Acs 2002, Lee et al. 2004). However, Sutaria, Hicks (2004) found that regions with larger firm size have faster firm birth rates.

The effect of unemployment on new firm formation is mixed. For instance, Reynolds et al. (1995) found a positive relationship between unemployment rate and firm births. Unemployed individuals can start new ventures in the earlier stages of unemployment. However, if unemployment persists for a long period, the propensity for self-employment declines (Ritsila, Tervo 2002). While Sutaria, Hicks (2004) found a negative relationship, Fritsch, Falck (2002) concluded that there is no relationship between new firm formation and unemployment. Storey (1991), Audretsch, Fritsch (1994), and Cheng, Li (2010) argued that depending on the type of estimation models, the unemployment rate can have both positive and negative association with new firm births.

Population growth also influences new start-ups. On the one hand, Audretsch, Fritsch

(1994), Davidsson et al. (1994), Guesnier (1994), Kangasharju (2000), and Armington, Acs (2002) noted that regional differences in firm birth rates can be explained by population growth. On the other hand, Sutaria, Hicks (2004) did not find any positive association between population growth and new firm formation and attributed the findings to modeling limitations.

Another predictor of new firm formation is per capita income growth. While Armington, Acs (2002) and Lee et al. (2004) found evidence of a positive effect, Sutaria, Hicks (2004) found no statistically significant effect of per capita income growth on new firm formation.

A number of studies have attested to the existence of spatial and temporal variation of entrepreneurship in the US. For instance, Armington, Acs (2002) suggested that entrepreneurial activities vary considerably across space, but not much over time. Acs, Mueller (2008) indicated that regional characteristics determine start-up rates. Urban, suburban, and rural differences in new business formation are also evident (Renski 2009). Parajuli, Haynes (2017) found that the spatio-temporal distributional patterns of single-unit firm births fluctuate in New England. In essence, whether within clusters, metropolitan areas, or states, new venture growth fluctuates over time and across space (see Guzman, Stern 2015, Morelix et al. 2016). Based on these findings, this study hypothesizes that the determinants of new firm formation influenced single-unit firm births within and across local areas (counties) in New England from 1999 through 2009 – a period covering the "great recession".

Since the effect of individual determinants on new firm formation varies, based on the general findings, it is hypothesized that regions with higher population densities should experience higher firm births. The availability of higher human capital and local financial capital should also be positively associated with new firm formation. However, new firm formation should be negatively associated with the size of the existing firms in terms of employment. As the unemployment rate does not have any consistent effect on new firm formation, it is hypothesized that the relationship between the unemployment rate and new firm births is indeterminate. Finally, higher population growth rate as well as higher per capita income growth rate should be positively associated with new firm formation.

3 Methodology and Data

Unlike pure cross-sectional and time series analyses, the panel data technique offers various advantages, such as ability to control for individual heterogeneity and allows for more variability while reducing the issue of multicollinearity and providing more degrees of freedom (see Baltagi 2005). However, it should be noted that the distribution of a variable of interest – new firm formation – often exhibits spatial heterogeneity and spatial autocorrelation when cross-sectional observations – counties – are spatial units (see Anselin 1988). As the non-spatial panel model will not be able to capture such effects and since spatial interaction effects in the form of spillovers are expected, it is necessary to calibrate spatial panel models that allow one to account for such effects in relation to the new firm formation dynamics.

For a panel of N observations over time T periods with K explanatory variables, a spatial panel regression model that includes spatial effects is (see Elhorst 2003, 2014):

$$Y_t = \rho W Y_t + \alpha \iota_N + X_t \beta + W X_t \theta + u_t \tag{1}$$

where u_t is the error component and is defined as:

$$u_t = \lambda W u_t + \epsilon_t \tag{2}$$

In equations (1) and (2), Y_t is an $NT \times 1$ vector of the dependent variable, ι_N is a $NT \times 1$ vector of the constant terms, X is an $NT \times K$ matrix of the independent variables, β and θ are both $K \times 1$ vector of the coefficients, and ϵ is a $NT \times 1$ vector of error terms that are independent and identically distributed with mean zero and variance σ^2. In addition, W is the spatial weight matrix of size $N \times N$, scalar ρ is the spatial autoregressive term, and scalar λ the spatial autocorrelation term.

The spatial weight matrix, W, describes the arrangement of possible interactions among spatial units. Such matrices could be based on the order of contiguity, inverse

distance, or nearest neighbors (see Elhorst 2014). For example, the queen contiguity weight matrix accounts for common edges and vertices of contiguous spatial units with respect to the reference spatial unit. In the case of a non-normalized, first-order queen weight matrix that takes binary values (0 and 1) weight coefficients assume 1 for the commonly shared edges and vertices in the immediate vicinity and 0 otherwise. The weight matrix is generally normalized for the ease of interpretation (see Elhorst 2014). As the spatial panel model takes into consideration spatial interactions based on the spatial weight matrix, in Equation (1), WY_t and WX_t represent the spatially weighted dependent and independent variable in the matrix form, respectively.

Based on the value of ρ, λ, and θ, different types of models can be specified:

- If $\lambda = 0$ and $\theta = 0$: Spatial autoregressive model (SAR)
- If $\rho = 0$ and $\theta = 0$: Spatial error model (SEM)
- If $\lambda = 0$: Spatial Durbin model (SDM)
- If $\lambda = 0$, $\rho = 0$, and $\theta = 0$: Non-spatial model

The global Moran's I statistic given by Equation (3) is used for examining spatial dependence (Bailey, Gatrell 1995).

$$I = \frac{N \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{N}(x_i - \bar{x})^2 (\sum_{i \neq j} \sum w_{ij})} \tag{3}$$

where w_{ij} is the element of weight matrix W, x is the variable of interest, and $\bar{x}$ is the mean of x.

There are a number of issues that have to be accounted for while calibrating a panel data model. First, the fixed effects model is generally preferred when there is a specific set of cross-sectional observations and the inference is based on them instead of the sample drawn from the population (see Baltagi 2005). Thus, as this study is focused only on New England, the fixed effects models will be adopted. Second, firm births in a county are likely to affect firm births in the neighboring/contiguous counties. That is, there are direct and indirect effects associated with new firm formation and these impacts can be considered in the SAR and SDM techniques. Thus, both SAR and SDM will be calibrated. Based on the log-likelihood and Akaike Information Criterion (AIC), the appropriate model will be selected. Third, as noted earlier, as new firms provide impetus to growth even during recessions (Schramm 2004), rates of new entrepreneurs have not abruptly changed during the observation period (Fairlie et al. 2016), and entrepreneurial activities change more across regions than over time (Armington, Acs 2002), only county-specific effects will be considered. The XSMLE module in Stata will be used for the estimation purpose (Belotti et al. 2016).

Data at the county level used in this study come from three different sources – the Census Bureau (CB), Bureau of Economic Analysis (BEA), and Bureau of Labor Statistics (BLS). The number of single-unit firm births, area (in square miles), public school enrollment, capital deposit (in US dollars) in local commercial and saving institutions, and establishment size in terms of employees are from the CB. Data from the BEA include per capital personal income (in US dollars), population, population growth rate, and per capita personal income growth. The unemployment rate was obtained from the BLS. Table 1 provides the details of variables.

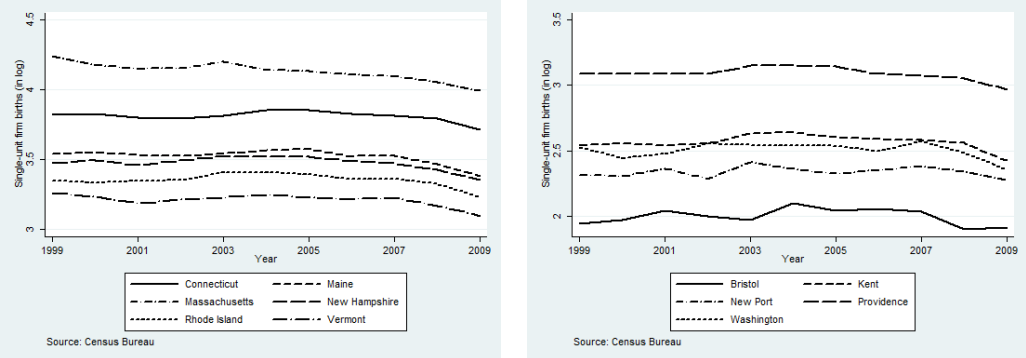
4 Empirical Results

The temporal variations of single-unit firm births (in logarithm) by states in New England and by counties in Rhode Island are shown in Figure 1.

In New England, over the study period, Massachusetts and Vermont experienced the largest and smallest number of single-unit firm births, respectively. In addition, at the state level, new firm formation was generally declining between 2007 and 2009. This

Table 1: Variable details

Variable	Description	Data Source
Dependent in logarithm		
lsub	Logarithm of single-unit firm birth	CB
Independent in logarithm		
lpopden	Logarithm of population density	BEA and CB
lpubenrpc	Logarithm of public enrollment per capita	BEA and CB
lfincappc	Logarithm of per capita deposit in local commercial and savings institutions	
lestratiol	Logarithm of the ratio of establishments with 50 or more employees to establishments with less than 10 employees	CB
luempr	Logarithm of unemployment rate	BLS
Independent in non-logarithm		
popgr	Population growth rate	BEA
perincgr	Per capital personal income growth rate	BEA



(a) State aggregate (b) County aggregate (Rhode Island)

Figure 1: Temporal variation of single-unit firm births

corresponded with the bottoming out of the great recession. As shown in Figure 1b, such a trend was also observable at the county level within a state.

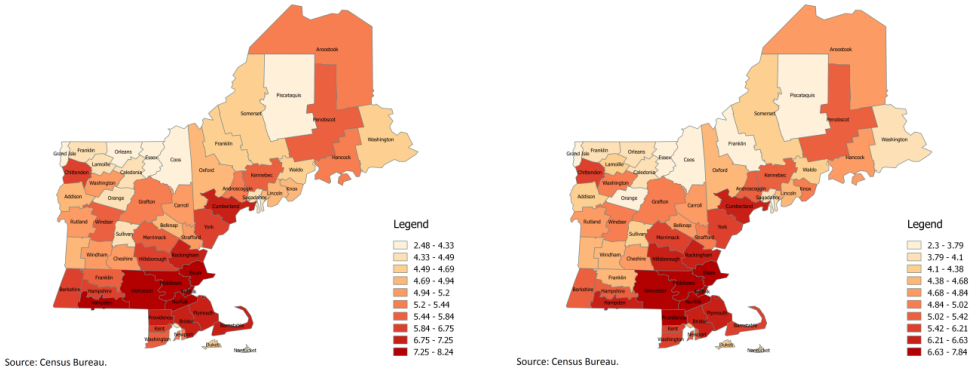
The spatial distributions of single-unit firm births (in logarithm) in 1999 and 2009 in New England are shown in Figure 2.

The estimated values of two non-spatial regression models – pooled and fixed effects panel – with the logarithm of single-unit firm births as the dependent variable are shown in Table 2.

In order to select the appropriate model between the pooled and fixed effects model, the F-statistic is used. The null hypothesis is that the intercept term (individual effects) is constant across all counties (i.e., the model is "poolable") and the alternative hypothesis is that the intercept term varies across counties (see Baltagi 2005). The F-statistic for degrees of freedom F(64, 643) is equal to 204.67 with p <0.000. This suggests that, at 5% significance level, the null is rejected and the fixed effects model is chosen. Further, log-likelihood and AIC values also suggest that the fixed effects model is appropriate compared to the pooled model.

As the unit of observations is not randomly selected, the fixed effects estimation makes sense. It should also be noted that the observation period is short and hence unit root analysis of individual time series is often not effective (see Baltagi 2005).

The non-spatial fixed-effects panel model suggests that for the 65 counties in New Eng-



(a) 1999 (b) 2009

Figure 2: Spatial variation of single-unit firm births in 1999 and 2009

land[1] over the period of observation (1999–2009), controlling for other factors, population density, public enrollment, and personal income growth have a statistically significant positive correlation with single-unit firm births. On the contrary, establishment size and unemployment rate have a significant negative association with new firm formation while holding other variables constant. County level financial capital and population growth rates were not significantly associated with firm births.

Before calibrating the spatial fixed-effects panel models, the Moran's I statistics for all variables across the study period are calculated. The global Moran's I statistics based on the first-order queen contiguity weight matrix are presented in Table 3.

Except for the human capital variable, that is, the logarithm of public school enrollment per capita, the global Moran's I statistic of variables are generally positive and statistically significant in each observation period. The results imply that each variable is autocorrelated in space and that spatial dependence exists.

Table 4 shows the estimated values of the fixed effects SAR and fixed effects SDM in which the logarithm of single-unit firm births is the dependent variable. Based on the log-likelihood and AIC values, the preferred model is the fixed effects SDM.

In the fixed effects SDM, the spatial dependence associated with single-unit firm births (ρ) is positive and statistically significant and suggests that new firm formation is spatially endogenous with respect to individual counties. Population density and public school enrollment per capita are both statistically significant and positively associated with single-unit firm births. Establishment size ratio is also statistically significant, but negatively associated with new firm formation.

Note that the SDM estimates are not interpreted as partial derivatives as in the classical regression technique. This is because a change in an explanatory variable in a county can impact single-unit firm births in other neighboring counties based on the spatial weight matrix (here, the queen contiguity matrix). Instead, the direct, indirect, and total effects as shown in Table 5 are interpreted (see LeSage, Pace 2009).

Population density has a positive direct and indirect impact on single-unit firm births. The difference between the coefficient estimate (0.945) and the direct effect estimate (1.109) is 0.164, which reflects a positive feedback to a county itself. That is, an increase in population density results in an increase in the number of new firms within a county. The spatially lagged coefficient of population density is not significant, but the indirect effect is positive and statistically significant and has a magnitude that is almost twice the magnitude of the direct effect. This suggests a large spillover resulting from the population density of a county to nearby counties. The total impact is 3.220 and statistically significant. Thus, ceteris paribus, a 10 percent increase in population density would result in more than a 32 percent increase in new firm formation and that about 1/3

[1] Although there are 67 counties in New England, this study only considers 65 contiguous counties.

Table 2: Non-spatial regression models

	Pooled	Fixed effects panel
Intercept	6.068***	
	(16.27)	
lpopden	0.397***	1.215**
	(20.72)	(3.48)
lpubenrpc	0.424*	0.304***
	(2.23)	(3.88)
lfincappc	0.234***	0.002
	(4.44)	(0.08)
lestratiol	0.839***	-0.315***
	(13.18)	(-5.25)
luempr	0.017	-0.266***
	(0.21)	(-8.18)
popgr	0.036	0.013
	(0.95)	(1.28)
perincgr	0.016†	0.007**
	(1.82)	(2.82)
Log-likelihood	-650.004	444.674
AIC	1316.008	-873.348
Observation	715	715

Source: Authors' calculations
Notes: Significant at † p<0.1, * p<0.05, ** p<0.01, ***p<0.001; t-values in parentheses

of this would have resulted from the direct impact and the remaining from the spillover effects.

Similarly, a 10 percent increase in public school enrollment per capita results in a 6 percent increase in single-unit firm births. About 33 percent of the increase in new firm formation results from the direct impact and 67 percent from the indirect impact of human capital. The significant total impact of the population growth rate on new firm formation is 0.082 and almost all of that impact is comprised of spillover effects. Thus, for a 1 unit change in the population growth rate, single-unit firm births are expected to increase by more than 8 percent ceteris paribus.

Total effects of the establishment size ratio and the unemployment rate are both negative and statistically significant. Thus, for a 10 percent increase in the size of existing firms, new firm formation decreases by 6.6 percent and for a 10 percent increase in the rate of unemployment, new firm formation decreases by 3.1 percent. In addition, in the case of the establishment size ratio, both the direct and indirect impacts influence

Table 3: Global Moran's I statistics

Year	lsub	lpopden	lpubenrpc	lfincappc	lestratiol	luempr	popgr	perincgr
1999	0.515***	0.764***	-0.024	0.095†	0.232**	0.396***	0.314***	0.341***
2000	0.504***	0.765***	-0.072	0.075	0.261***	0.484***	0.356***	0.390***
2001	0.522***	0.764***	-0.067	0.079†	0.254***	0.360***	0.351***	0.178*
2002	0.522***	0.765***	-0.098	0.106†	0.244**	0.287***	0.302***	0.143*
2003	0.521***	0.765***	-0.118	0.137*	0.242**	0.412***	0.244**	0.217**
2004	0.524***	0.765***	-0.105	0.146*	0.216**	0.559***	0.289***	-0.003
2005	0.522***	0.764***	-0.091	0.118**	0.219**	0.610***	0.205**	0.169*
2006	0.541***	0.764***	-0.093	0.144**	0.242**	0.547***	0.055	0.206**
2007	0.524***	0.764***	-0.047	0.139**	0.245**	0.503***	0.051	0.024
2008	0.510***	0.764***	-0.097	0.141**	0.189**	0.569***	0.157*	0.021
2009	0.540***	0.763***	-0.078	0.052	0.207**	0.513***	0.147*	0.365***

Source: Authors' calculations
Notes: Significant at † p<0.1, * p<0.05, ** p<0.01, ***p<0.001

Table 4: Spatial panel models

	Fixed effects SAR	Fixed effects SDM
lpopden	0.781**	0.945**
	(2.77)	(2.98)
lpubenrpc	0.200**	0.146*
	(3.16)	(2.25)
lfincappc	-0.001	0.019
	(-0.04)	(0.83)
lestratiol	-0.245***	-0.260***
	(-5.06)	(-5.37)
luempr	-0.138***	-0.036
	(-5.04)	(-0.75)
popgr	-0.001	-0.006
	(0.16)	(-0.71)
perincgr	0.003†	0.002
	(1.67)	(1.02)
Lagged lsub (ρ)	0.532***	0.497***
	(15.65)	(13.40)
W*lpopden		0.668
		(1.33)
W*lpubenrpc		0.160
		(1.35)
W*lfincappc		-0.016
		(-0.33)
W*lestratiol		-0.075
		(-0.73)
W*luempr		-0.117*
		(-1.97)
W*popgr		0.046**
		(3.20)
W*perincgr		0.002
		(0.60)
Log-likelihood	536.345	547.447
AIC	-1054.689	-1062.894
Observations	715	715

Source: Authors' calculations
Notes: Significant at † p<0.1, * p<0.05, ** p<0.01, ***p<0.001; t-values in parentheses.

single-unit firm births, while in the case of the unemployment rate; only the spillover effect is significantly associated with single-unit firm births.

Finally, there was no significant relationship or spillover effect associated with the availability of local financial capital and personal income growth rates with respect to new firm formation.

5 Conclusion

Entrepreneurial ventures are important for the US economy, and this paper examined the association between new firm formation and its determinants as well as the spillover effects in New England at the county level from 1999 through 2009 using both non-spatial and spatial panel data models. Based on the SDM, it was found that population density had the largest, positive impact on single-unit firm births both within a county and among its neighbors and that the spatial spillover was larger than the comparable direct effect on a county. This suggests that agglomeration and urbanization are conducive to new firm formation and that rural areas are less favorable for starting new businesses. Human capital – formal education, professional and managerial skills, information training, and

Table 5: Estimated direct, indirect, and total effects

	Direct	Indirect	Total
lpopden	1.109***	2.111**	3.220***
	(4.05)	(2.62)	(3.56)
lpubenrpc	0.184*	0.416*	0.600*
	(2.44)	(1.97)	(2.41)
lfincappc	0.019	-0.010	0.009
	(0.69)	(-0.10)	(0.08)
lestratiol	-0.290***	-0.372*	-0.662**
	(-5.88)	(-2.01)	(-3.18)
luempr	-0.048	-0.264***	-0.312***
	(-1.04)	(-3.97)	(-4.25)
popgr	0.002	0.080**	0.082**
	(0.26)	(3.19)	(2.76)
perincgr	0.002	0.004	0.006
	(1.00)	(0.77)	(1.14)

Source: Authors' calculations
Notes: Significant at † p<0.1, * p<0.05, ** p<0.01, ***p<0.001; t-values in parentheses.

innovative abilities – is equally crucial for creating new ventures. A region with a high level of human capital will not only foster new firms within its boundary, but will also influence new firm formation in the neighboring areas. In addition, population growth is important in creating new start-ups both locally as well as in the nearby areas. This could be attributed to the fact that as the number of individuals increases, so does the demand for services and the potential entrepreneurs. A variety of earlier research findings support the results that population density, human capital, and population growth influence new firm formation (for example Fritsch 1992, Audretsch, Fritsch 1994, Reynolds et al. 1994, Armington, Acs 2002).

On the contrary, individuals are more likely to avoid the risk and uncertainties associated with new businesses and instead join existing firms that provide alternative employment and income-generation opportunities when such opportunities are available. Thus, as alternative establishment size increases, new firm formation decreases and such trends produce significant negative spillovers in neighboring areas. Similarly, as the rate of unemployment increases, single-unit firm births decrease. With increasing unemployment, individuals might migrate to other regions seeking opportunities or could rely on government welfare instead of starting opportunity or necessity based entrepreneurial ventures. Some earlier studies have found a significant negative association between new firm formation and establishment size as well as between firm births and unemployment (for example Armington, Acs 2002, Sutaria, Hicks 2004). Finally and unexpectedly, this study did not find any significant association between the availability of local financial capital or personal income growth with single-unit firm births.

A region endowed with various determinants of entrepreneurial ventures will not only influence itself, but its neighboring regions. Thus, rather than merely promoting policies to attract entrepreneurial ventures for economic growth and development, local/regional policymakers need to understand the nature of spatial externalities associated with new firm formation and should focus on how to benefit from them.

This study calibrated (non)spatial models using 65 contiguous counties and dropped two island counties. Future research should use other types of weighting schemes, such as distance- and nearest neighbor-based weight matrices, for the estimation purpose. However, it is hypothesized that the results would be similar. Moreover, as the focus was only on New England, future work should examine other regions of the US and may want to examine more explicitly the temporal considerations. While this study used the aggregate number of single-unit firm births, by disaggregating new firm formation by various industry sectors and calibrating sector-based panel models new insights related to economic structure might be provided.

Acknowledgement

The authors would like to thank Roger R. Stough for the data on new firm formation. They are also grateful to the editors and anonymous reviewer for the constructive feedback.

References

Acs ZJ, Armington C (2004) The impact of geographic differences in human capital on service firm formation rates. *Journal of Urban Economics* 56: 244–278. CrossRef.

Acs ZJ, Mueller P (2008) Employment effects of business dynamics: Mice, gazelles and elephants. *Small Business Economics* 30: 85–100. CrossRef.

Acs ZJ, Szerb L, Autio E (2015) Global entrepreneurship index 2016. The Global Entrepreneurship and Development Institute, Washington, DC. https://www.researchgate.net/publication/284727510_Global_Entrepreneurship_Index_2016. CrossRef.

Anselin L (1988) *Spatial econometrics: Methods and models.* Kluwer, Dordrecht, The Netherlands. CrossRef.

Anyadike-Danes M, O'Reilly M (2005) Watch that space! The county hierarchy in firm births and deaths in the UK, 1980-1999. *Small Business Economics* 25: 273–292. CrossRef.

Armington C, Acs ZJ (2002) The determinants of regional variation in new firm formation. *Regional Studies* 36: 33–45. CrossRef.

Audretsch DB, Fritsch M (1994) The geography of firm births in Germany. *Regional Studies* 28: 359–365. CrossRef.

Bailey TC, Gatrell AC (1995) *Interactive spatial data analysis.* Pearson, Essex, UK

Baltagi BH (2005) *Econometric analysis of panel data.* John Wiley & Sons, West Sussex, UK

Belotti F, Hughes G, Mortari AP (2016) Xsmle: Stata module for spatial panel data models estimation. https://ideas.repec.org/c/boc/bocode/s457610.html

Bosma N, Schutjens V (2011) Understanding regional variation in entrepreneurial activity and entrepreneurial attitude in Europe. *Annals of Regional Science* 47: 711–742. CrossRef.

Cheng S, Li H (2010) The effects of unemployment on new firm formation revisited: Does space matter? *Regional Science Policy & Practice* 2: 97–120

Dahl MS, Sorenson O (2012) Home sweet home: Entrepreneurs' location choices and the performance of their ventures. *Management Science* 58: 1059–1071. CrossRef.

Davidsson P, Honig B (2003) The role of social and human capital among nascent entrepreneurs. *Journal of Business Venturing* 18: 301–331. CrossRef.

Davidsson P, Lindmark L, Olofsson C (1994) New firm formation and regional development in Sweden. *Regional Studies* 28: 395–410. CrossRef.

Delfmann H, Koster S, McCann P, van Dijk J (2014) Population change and new firm formation in urban and rural regions. *Regional Studies* 48: 1034–1050. CrossRef.

Delgado M, Porter ME, Stern S (2010) Clusters and entrepreneurship. *Journal of Economic Geography* 10: 495–518. CrossRef.

Elhorst JP (2003) Specification and estimation of spatial panel data models. *International Regional Science Review* 26: 244–268. CrossRef.

Elhorst JP (2014) Spatial panel models. In: Fischer MM, Nijkamp P (eds), *Handbook of Regional Science*. Springer-Verlag, Heidelberg, Germany, 1637–1652. CrossRef.

Ellison G, Glaeser EL (1999) The geographic concentration of industry: Does natural advantage explain agglomeration? *American Economic Review* 89: 311–316. CrossRef.

Fairlie RW, Morelix A, Reedy EJ, Russell J (2016) The Kauffman index: Startup activity national trends. Ewing Marion Kauffman Foundation, Kansas City, MO, US, http://www.kauffman.org/~/media/kauffman_org/microsites/kauffman_index/-startup_activity_2016/kauffman_index_startup_activity_national_trends_2016.pdf

Fritsch M (1992) Regional differences in new firm formation: Evidence from West Germany. *Regional Studies* 26: 233–241. CrossRef.

Fritsch M, Falck O (2002) New firm formation by industry over space and time: A multi-level analysis. Working Paper No. 11, Freiberg University of Mining and Technology, Freiberg, Germany. http://www.wiwi.uni-jena.de/uiw/publications/pub_1999_2003/fritsch_falck_2002.pdf

Garvin DA (1983) Spin-offs and the new firm formation process. *California Management Review* 25: 3–20. CrossRef.

Guesnier B (1994) Regional variations in new firm formation in France. *Regional Studies* 28: 347–358. CrossRef.

Guzman J, Stern S (2015) Where is silicon valley? *Science* 347: 606–609

Kangasharju A (2000) Regional variations in firm formation: Panel and cross-section data evidence from Finland. *Papers in Regional Science* 79: 355–373. CrossRef.

Kim PH, Aldrich HE, Keister LA (2006) Access (not) denied: The impact of financial, human, and cultural capital on entrepreneurial entry in the United States. *Small Business Economics* 27: 5–22. CrossRef.

Krugman P (1991) Increasing returns and economic geography. *Journal of Political Economy* 99: 483–499. CrossRef.

Lee SY, Florida R, Acs ZJ (2004) Creativity and entrepreneurship: A regional analysis of new firm formation. *Regional Studies* 38: 879–891. CrossRef.

LeSage J, Pace RK (2009) *Introduction to spatial econometrics*. CRC Press, Boca Raton, FL, US. CrossRef.

Michelacci C, Silva O (2007) Why so many local entrepreneurs. *Review of Economics and Statistics* 89: 615–633. CrossRef.

Morelix A, Reedy EJ, Fairlie RW, Russell J (2016) The Kauffman index: Startup activity state trends. Ewing Marion Kauffman Foundation, Kansas City, MO, US. http://www.kauffman.org/~/media/kauffman_org/microsites/kauffman_index/-startup_activity_2016/kauffman_index_startup_activity_state_trends_2016.pdf

NBER (2009) US business cycle expansions and contractions. http://www.nber.org/-cycles/cyclesmain.html

Parajuli J, Haynes KE (2017) An exploratory analysis of new firm formation in New England. In: Lombard J, Stern E, Clarke G (eds), *Applied Spatial Modelling and Planning*. Routledge, Abingdon, UK, 17–32

Renski H (2009) New firm entry, survival, and growth in the United States. *Journal of the American Planning Association* 75: 60–77. CrossRef.

Reynolds PD, Miller B, Maki WR (1995) Explaining regional variation in business births and deaths: U.S. 1976-88. *Small Business Economics* 7: 389–407. CrossRef.

Reynolds PD, Storey DJ, Westhead P (1994) Cross-national comparisons of the variation in new firm formation rates. *Regional Studies* 28: 443–456. CrossRef.

Ritsila J, Tervo H (2002) Effects of unemployment on new firm formation: Micro-level panel data evidence from Finland. *Small Business Economics* 19: 31–40. CrossRef.

Robinson MJ, Cottrell TJ (2007) Investment patterns of informal investors in the Alberta private equity market. *Journal of Small Business Management* 45: 47–67. CrossRef.

Robinson PB, Sexton EA (1994) The effect of education and experience on self-employment success. *Journal of Business Venturing* 9: 141–156. CrossRef.

Schramm CJ (2004) Building entrepreneurial economies. *Foreign Affairs* 83: 104–115. CrossRef.

Storey DJ (1991) The birth of new firms: Does unemployment matter? A review of the evidence. *Small Business Economics* 3: 167–178. CrossRef.

Sutaria V, Hicks DA (2004) New firm formation: Dynamics and determinants. *Annals of Regional Science* 38: 241–262. CrossRef.

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 3, 2017, 77–86
DOI: 10.18335/region.v4i3.169

journal homepage: region.ersa.org

Urban Concentration and Spatial Allocation of Rents from natural resources. A Zipf's Curve Approach

Tomaz Ponce Dentinho[1]

[1] University of Azores, Azores, Portugal

Received: 17 October 2016/Accepted: 18 May 2017

Abstract. This paper aims at demonstrating how countries' dependency on natural resources plays a crucial role in urban concentration. The Zipf curve elasticity is estimated for a group of countries and related to a set of indicators of unilateral transferences. Results show that in comparison to others, countries with higher urban concentration explained by higher Zipf curve elasticity have a higher percentage of income coming from natural resources and education expenditures whereas public spending in health and net outflow of foreign direct investment seem to have spatial redistribution effects. Summing up, there are signs that the spatial allocation of property rights over natural resources and related rents influences urban concentration.

Key words: Urban concentration; Zipf Curve; Natural Resources; Rents

1 Introduction

Globally, more people live in urban areas than in rural areas and the world's fastest growing cities are located in Africa and Asia (United Nations 2014), these being increasingly connected to the complex constellation of the "New Urban World" (Kourtit, Nijkamp 2013). Notwithstanding this, urban concentration in the developing world is not necessarily a good thing and neither is it unmanageable. On the one hand, urban concentration can generate congestion, environmental disturbances and social problems that are difficult to address with reduced resources. On the other hand, urban concentration does not happen only due to global and undifferentiated factors, but may relate to manageable causes such as unilateral transferences between regions.

The 'New Urban World' is certainly interacting through competition, collaboration, exploitation and help. Actually, unilateral transferences between regions can occur when the rents from natural resources, distributed across the territory, go to the regions and places where their owners receive them. Permanent governmental transferences happen when taxes collected across the territory go to capital cities and to border areas where public services tend to be relatively more concentrated. Migrant remittances persist throughout time between immigrant places and emigrant localities. This is part of the migrant strategy that transfers a portion of their income to the areas where their families live. Finally, investment flows may lead also to spatial distribution of income. All of these unilateral and quite enduring transferences create persistent multiplier effects that accumulate in an uneven concentration of production and expenditure, employment and population, tradable and non-tradable activities.

Unilateral transferences between regions and the multiplier income effects associated with them explain why the spatial distribution of production may not be the same as the spatial distribution of expenditure. Therefore, it is not accurate to look at local and regional development and to the concomitant effects of urban concentration and rural desertification without taking into account unilateral transferences between places and regions.

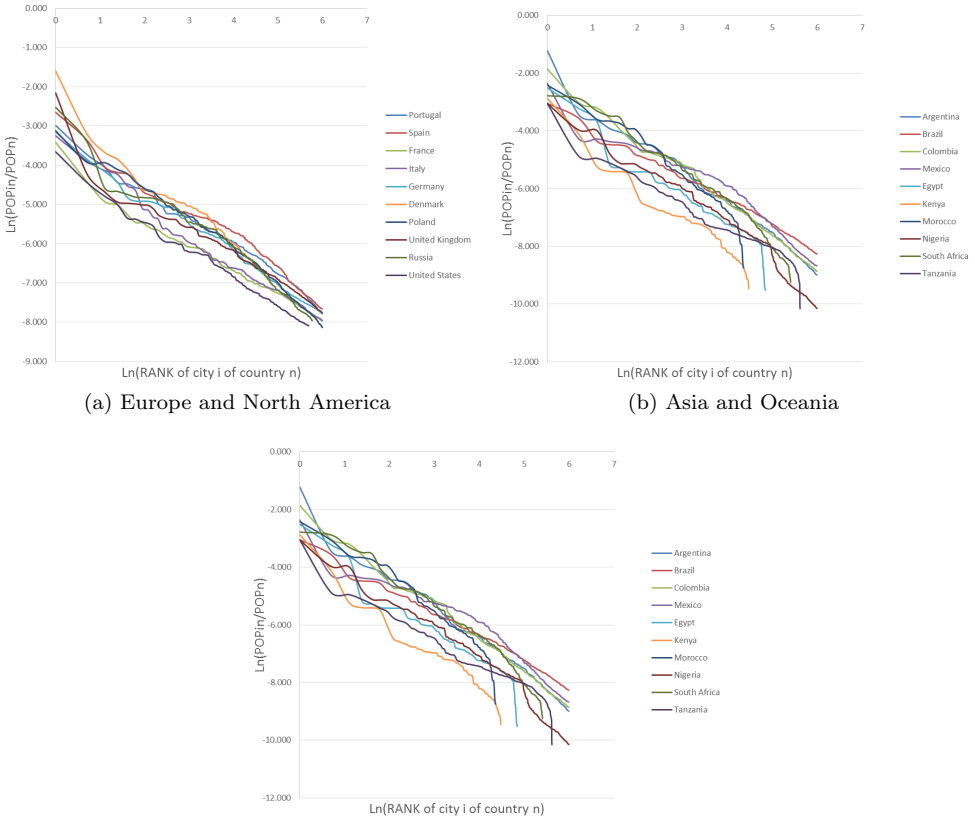
In fact, the spatial allocation of rents of natural resources depends on the spatial distribution of property rights over natural resources and not on the territorial distribution of those resources. Due to the spatial distribution of property rights and the distribution of territorial resources being different, it is quite likely that permanent unilateral transferences occur and that dependent and exploited regions are created and maintained. Dependent regions are those, usually in the centre or at the territorial borders, who receive permanent transferences to obtain and secure the rents of territorial property rights. Exploited regions are those deprived from the rents of their territory (Dentinho 2012, 2017). The isolated estate of Von Thünen (1826) is a crucial reference if one wants to understand this, given that the author highlighted the creation of land rents, leaving the rent distribution implicit in the model. In the Von Thünen model, land rents go from territorial managers located in peripheral rings to the land owners located in the city centre; there the received rents generate income multiplier effects as thought by François Quesnay (1758). In the end, land ownership and rent distribution constitute crucial elements for urban concentration, regional development and, with free migration.

In extreme cases, one whole country can be concentrated into just one major city receiving the rents from a territory with much smaller localities. But in theory and in reality the hierarchy of cities is quite resilient as it is showed by Gibrat's Law (1931) and by the Zipf law (1949), which states that the second largest city is half the size of the biggest, the third largest city is a third the size of the biggest, and so on. As said by Gabaix (1999) this means that the probability of the size of a city being greater than some limit S is proportional to $1/S$: P (size$>$ S)$= C/S^{\mu}$; where $\mu = 1$ and $C =$ Constant.

There are many applications of Zipf's estimates focusing on the population of cities, dimension of companies, webpage popularity, impact of tornados and earthquakes as well as a few other phenomena (Pinto et al. 2012). In an application to the Chinese cities (Anderson, Ying 2005) it is concluded that cities tend to stay in the same position in the ranking and that the lognormal distribution explains the ranking profile better than the Pareto distribution. Jiang et al. (2015) tested the stability of the Pareto coefficient for truncated series of cities in China. Black, Henderson (2003) work on United States Cities found out that there is a concentration in service cities. Bosker et al. (2008) looked at the evolution of German cities from 1925 until 1999 and proved that, besides the enormous shocks caused by the economic crises of the thirties, world war II, the cold war and the fall of the Berlin Wall, there are increasing returns on scale for city growth. In an application to South Africa (Morudu 2016) it was suggested that the Zipf curve could be used to estimate the population of cities for a determined year where census data is not available.

It is also common to find in the literature theoretical analysis of the Zipf curve. Cristelli et al. (2012) demonstrate that truncated series of city population influences the estimates of Zipf coefficients. Along with the work of Anderson, Ying (2005), Benguigui, Blumenfeld-Lieberthal (2007) argue that the Zipf curve can be concave, linear or convex depending on the country. Peng (2010)Peng (2010) proves that the shape of the Zipf curve relates to the development of the network of cities in China. Cordoba (2008) shows that the profile of the Zipf curve comes from the regularity of the growth process across cities. Many authors test the validity of different distributions to explain the population size of the cities (Loannides, Overman 2003, Nitsch 2005, Newman 2005, Giesen et al. 2010, Gómez-Déniz et al. 2014, Gómez-Déniz, Calderín-Ojeda 2015, Shujuan 2016, Luckstead, Devadoss 2017).

Finally, there are many attempts to explain the regularity of the hierarchy of cities. Ades, Glaeser (1995) found that political factors do influence urban concentration and that dictatorships occur in central cities that are, on average, 50% larger than democratic countries of similar size. Krugman (1996) suggests that nature is also hierarchical in river basins and that cities rooted in natural capital have also a string hierarchy. Eaton,



(a) Europe and North America

(b) Asia and Oceania

(c) Africa and South America

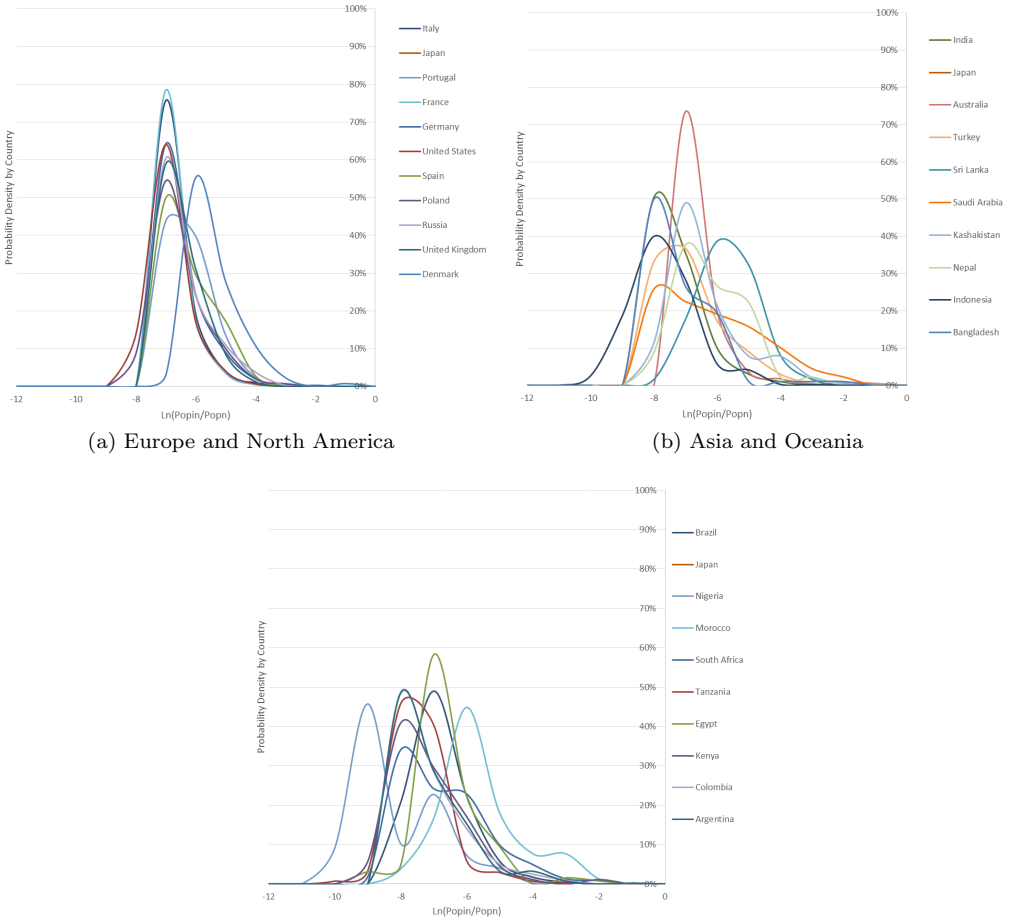
Figure 1: Zipf 's curve by Country

Eckstein (1997) relate the hierarchy to intercity migration. Gabaix (1999) connects the Zipf law with relative city growth and points out that parameter $\mu = 1$ does not change much with the age of the city network. Axtell, Florida (2001) reconcile the tension between centripetal and centrifugal forces. Duranton (2002) links the city hierarchy with a set of indicators related to innovation. Córdoba (2003) shows that the Zipf law happens when there is a similar growth path between cities. Giesen, Suedekum (2014) find that older cities tend to be larger than new ones. There is interesting literature linking urban concentration to economic growth (Bertinelli, Strobl 2007) where there is an optimal level of urban concentration that can be influenced by policy makers (Henderson 2003, Brülhart, Sbergami 2009). Kourtit, Nijkamp (2013) report that recently major cities, beyond being the engine for growth, show a rapid population rise while smaller ones shrink and can even fall below a critical sustainable population level. To my knowledge, there is not an analysis relating urban concentration to the distribution of rents from natural resources and other unilateral transfers although they very much reveal the institutional context of the territory that Ades, Glaeser (1995) tried to relate to the hierarchy of cities.

The aim of this paper is to look into urban concentration to test the hypothesis that the spatial allocation of rents from natural resources and other unilateral transferences do have a role in urban concentration. Section 2 presents the estimates of the Zipf elasticity for various countries. Section 3 links the estimated elasticity with the features of those countries and to unilateral transferences. Finally in Section 4 the discussion is based on the relation between spatial allocation of property rights and urban concentration.

2 Zipf's Law and the level of Urban Concentration

Data of city population by country was obtained from the site http://worldpopulationreview.com/countries/. The population of each city (Pop_i) was divided by the total



(a) Europe and North America

(b) Asia and Oceania

(c) Africa and South America

Figure 2: Kernel Density of the Logarithm of the Relative Size of the Cities by Country

population of the country (Pop) obtaining the weight of each city in the total population of the country. Then the coefficient of the Zipf curves by Country (μ) (Figure 1) was estimated regressing the logarithm of the weight of the city population with the logarithm of the rank order of the city (Equation 1).

$$\ln\left(\frac{Pop_i}{Pop}\right) = C - \mu r_i \qquad \text{for } n \text{ countries} \tag{1}$$

The Zipf curves for European and North American countries show, with the exception of Russia, lower slopes than in other parts of the world. Nevertheless, there are countries, like the United Kingdom or Denmark, where the major city reduces the importance of the following ones. In other countries, a set of major cities reduces the weight of those who are next in the hierarchy, this being the case of Russia and Spain. Finally, the middle part of the curves do not show a very monotonous profile, which may indicate the existence of unilateral transferences also involving middle ranked cities, with this appearing to be more relevant in the United States and Italy. The Zipf curves for Asian and Oceania countries show tails in the distribution of cities indicating that parts of the countries are very depopulated due to mountain areas (Pakistan and Nepal), ocean archipelagos (Japan and Indonesia) or by long lasting conflicts (Sri Lanka). On the other hand, research rich countries have higher concentrations of their resource rents in major cities (Saudi Arabia, Kazakhstan). The Zipf curves for Africa and South American countries also show tails associated to remote areas, mainly in African countries (Kenya, Morocco, Egypt, Nigeria, South Africa and Tanzania), and also in countries from South America (Argentina, Brazil, Colombia and Mexico) with a larger concentration of population in major towns.

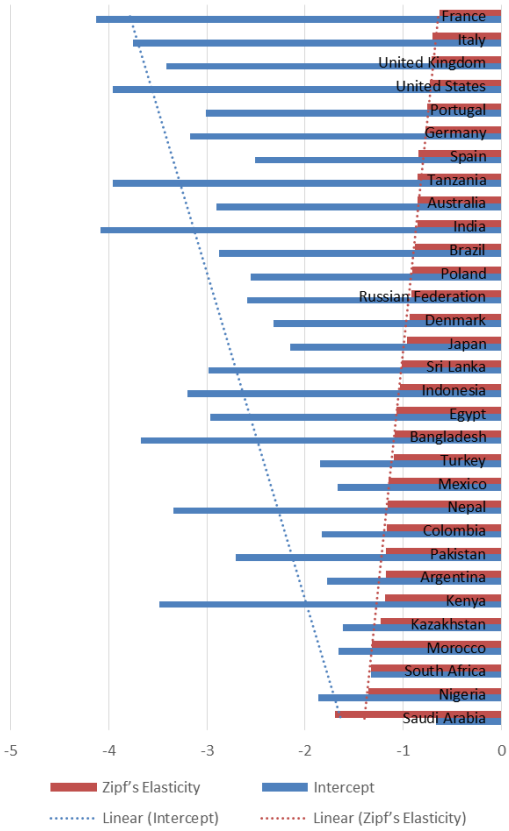


Figure 3: Zipf 's elasticity and intercept by country

Figure 2 shows the kernel densities of the logarithm of the relative size of the cities by country. Looking closer it is possible to identify various groups of countries. The countries in Europe and North America present very similar Pareto type distributions with more cities in the lower ranks of the distribution. For smaller countries close to bigger ones, like Denmark and somehow Portugal, there is a shift to smaller cities in the distribution. The distributions for countries in Asia, Oceania, Africa and South America are quite different from each other. Some Asian countries like India, Bangladesh and Australia behave like European and North American countries. A second group of countries do not have a mode in the cities with lower size; this happens with countries that have some remote and small cities like Brazil, the United States, Japan, Morocco, South Africa, Tanzania, Sri Lanka, Kazakhstan, Pakistan, Nepal, Poland and Indonesia. Finally, there are countries with two modes in the distribution of city sizes showing that the first cities emptied the second ones; this is the case of Nigeria, Turkey, Egypt, Kenya, Colombia, Mexico and Argentina.

The estimates of the Zipf curves by country are presented in Figure 3. As expected the intercept (C) increases when the Zipf coefficient (μ) decreases. Countries that do not match this expectable rule are the ones that have a reduced sample of cities: Tanzania, India, Bangladesh, Nepal, Kenya and Saudi Arabia.

Results in Figure 2 show that, although the value expected for the Zipf coefficient in the literature ($\mu = 1$) is the average estimate, there are quite significant differences between more concentrated countries like Saudi Arabia ($\mu = -1.698$) and more decentralized countries such as France ($\mu = -0.634$).

If people can migrate from less affluent to more affluent places, the expected value of the Zipf elasticity is around 1, as proved by (Gabaix 1999). For values of the Zipf elasticity higher than 1 there is a concentration of the population in the bigger cities whereas for values lower than that threshold there is a distribution of the population in the smaller cities. This paper attempts to find out if the spatial allocation of rents from natural resources and from other unilateral regional transferences (private and public)

Table 1: Estimated Zipf curves coefficients and data on unilateral transferences (World Bank 2010)

Country name	Number of cities	Intercept	Zipf's elasticity	Total natural resources rents[1]	Public health expenditure[1]	Adjusted savings: Education expenditure[1]	FDI, net outflows[1]	Workers' remittances[1]
Saudi Arabia	89	-0.671***	-1.698***	53.73	2.7	7.186	0.867	0.05
Nigeria	400	-1.868***	-1.352***	32.56	1.92	0.850	0.465	5.1
South Africa	223	-1.327***	-1.326***	4.62	3.94	5.434	-0.044	0.31
Morocco	78	-1.661***	-1.317***	2.61	1.97	5.201	0.639	7.07
Kazakhstan	192	-1.614***	-1.230***	27.60	2.55	4.413	5.294	0.2
Kenya	88	-3.482***	-1.182***	1.33	2.11	5.916	0.005	5.52
Argentina	400	-1.778***	-1.178***	6.07	4.42	6.010	0.262	0.17
Pakistan	400	-2.703***	-1.173***	3.92	0.84	1.565	0.027	5.48
Colombia	400	-1.833***	-1.165***	7.86	5.52	3.907	2.272	1.41
Nepal	45	-3.341***	-1.162***	3.39	1.83	4.202	0.000	21.66
Mexico	400	-1.664***	-1.147***	7.29	3.09	4.798	1.310	2.13
Turkey	400	-1.846***	-1.099***	0.45	5.07	2.644	0.200	0.12
Bangladesh	98	-3.673***	-1.089***	3.28	1.17	1.811	0.000	10.81
Egypt	127	-2.968***	-1.073***	10.13	1.74	4.411	0.537	3.53
Indonesia	385	-3.202***	-1.037***	5.95	1.28	4.328	0.376	0.98
Sri Lanka	62	-2.984***	-1.017***	0.50	1.32	1.744	0.086	8.38
Japan	514	-2.149***	-0.962***	0.03	7.83	3.185	1043	0.03
Denmark	130	-2.317***	-0.937***	2.13	9.71	7.278	1.06	0.2
Russian Fed.	319	-2.591***	-0.914***	19.90	3.15	3.544	3.531	0.35
Poland	400	-2.557***	-0.907***	0.98	5.42	4.869	1.177	1.62
Brazil	400	-2.876***	-0.882***	5.30	4.24	5.162	0.541	0.19
India	200	-4.086***	-0.857***	4.04	1.18	3.082	0.781	3.21
Australia	400	-2.900***	-0.852***	8.31	5.93	4.536	2.167	0.43
Tanzania	275	-3.956***	-0.852***	6.78	4.05	2.393	0.000	0.11
Spain	400	-2.513***	-0.842***	0.06	6.95	4.210	2.710	0.76
Germany	400	-3.175***	-0.771***	0.12	8.97	4.394	3.392	0.35
Portugal	355	-3.013***	-0.759***	0.33	7.50	4.937	-3.211	1.56
USA	295	-3.954***	-0.730***	1.00	9.49	4.789	2.432	0.04
UK	400	-3.413***	-0.712***	1.49	8.08	5.054	1.320	0.33
Italy	400	-3.755***	-0.699***	0.17	7.4	4.411	1.595	0.33
France	400	-4.127***	-0.634***	0.07	9.25	5.053	3.311	0.61

Notes: [1]In percent of GNI

has a role in the urban concentration profile represented by the Zipf coefficient.

3 Zipf's Law and the level of Urban Concentration

To answer this question the positive value of the Zipf elasticity for each country is regressed against different variables that relate to unilateral income transferences (Data in Table 1):

1. The Total natural resources rents as a percentage of gross national income;
2. Public heath expenditure as a percentage of gross national income.
3. Adjusted savings: education expenditure as a percentage of gross national income.
4. Net outflows of foreign direct investment as a percentage of gross national income;
5. Workers remittances also as a percentage of gross national income;

The gross national income per capita was included in the model and provided significant results as indicated by Pinto et al. (2012). Nevertheless, there were problems of colinearity with the public health expenditure and it is more informative for the present exercise

Table 2: Correlation between dependent and independent variables

	Zipf's elasticity	Total natural resources rents	Public health expenditure	Adjusted savings: Education expenditure[2]	FDI, net outflows	Workers' remittances
Zipf's elasticity	1.00	-0.65	0.64	-0.05	0.18	-0.26
Total natural resources rents	-0.65	1.00	-0.35	0.08	0.19	-0.10
Public health expenditure	0.64	-0.35	1.00	0.38	0.28	-0.51
Adjusted savings: Education expenditure (% of GNI)	-0.05	0.08	0.38	1.00	0.10	-0.30
FDI, net outflows	0.18	0.19	0.28	0.10	1.00	-0.32
Workers' remittances	-0.26	-0.10	-0.51	-0.30	-0.32	1.00

Notes: [2]In percent of GNI

Table 3: Linear regression to explain the Zipf curve elasticity by country

	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.
Adj R^2		0.406		0.591		0.628		0.634		0.629
F		21.51		22.67		17.88		14.02		11.19
Intercept	0.923	0.000	1.121	0.000	1.014	0.000	1.017	0.000	0.967	0.000
Total natural resources rents[3]	0.013	0.000	0.010	0.001	0.009	0.002	0.010	0.001	0.010	0.001
Public Health expenditure[3]			-0.038	0.001	-0.048	0.000	-0.042	0.001	-0.037	0.010
Adjusted savings: Education expenditure[3]					0.037	0.062	0.035	0.079	0.036	0.074
FDI net outflows[3]							-0.023	0.234	-0.021	0.280
Workers' remittances[3]									0.008	0.428

Notes: [3]In percent of GNI

to relate urban concentration with the spatial distribution of public spending than with the income level of the country. Table 2 presents the correlation between dependent and independent variables that indicates the strong positive correlation between the Zipf elasticity and the total natural resources rents and the strong negative correlation between the Zipf elasticity and the public health expenditures also slightly correlated with Workers' remittances.

Table 3 presents the estimated parameters of linear models that relate the urban concentration by country assessed by the Zipf elasticity with a set of unilateral transferences by country. Results are quite interesting:

First, the Percentage on the GNI of the Total Natural Resource Rents has a positive, significant and consistent relation with the Zipf city elasticity by country. Each percent point of the Total natural resources rents on the GDP increases the Zipf city elasticity by around 0.010. In other words, assuming that natural resources are distributed across the territory, bigger cities seem to own the property rights over natural resources across the territory and, through the multiplier effects of those rents, it increases their weight in the country population. That is why countries like Saudi Arabia, Nigeria and Kazakhstan – with higher percentage of the rents of natural resources on the total GNI – attract more population to the bigger city.

Second, the percentage of public health expenditure in the gross national income has a spatial distributive effect that is consistent across the various regressions with a quite robust coefficient that varies around 0.038. Interestingly the percentage of the "Adjusted savings: Education expenditure" in the gross national income has an urban concentration effect of the same scale on the distributive effect of public health spending.

Third, although not significant, countries that are net exporters of foreign direct investment tend to have less urban concentration, indicating somehow that investment

that goes outside goes also inside the country.

Finally, the Percentage on the GNI of the Workers Remittances has no robust effect in the urban concentration of the country. Furthermore as shown in Table 2, this is very much correlated with public health expenditures limiting the capacity of the results of Model 5. The attempt was to know if there was an indication that migrants in other countries tend either to secure population in smaller cities by sending remittances to their families that reside in those areas, or to invest their savings in more urbanized areas. A more profound analysis of differences between countries could be analyzed.

4 Discussion

The hypothesis tested in this exercise is quite simple: the spatial allocation of property rights over territorial resources has strong effects on the spatial profile of the urban network through the multiplier effects of income associated with rents from natural resources that goes from the places where those resources are located to the places of residence of their owners. The test relates the urban concentration of each country, assessed by the elasticity of the respective Zipf curve, with the percentage of income coming from the rents of natural resources.

Results confirm what Ades, Glaeser (1995) said: Institutions, namely those created by the spatial distribution of property rights over natural resources and by the geographical allocation of public spending, do have a role in the urban concentration throughout space. Furthermore, since huge amounts of rents from natural resources come from land, it is obvious, also from the research presented in this essay, that the main driver of major cities is Real Estate as justified by Castells (2012). To some extend we confirm the old intuition of the Von Thünen model that assumed that land rents go from territorial managers located in peripheral rings to the landowners located in the city centre where rents generate income multiplier effects as envisaged by Quesnay (1758). Regarding the Duranton (2002) argument, namely that urban concentration and innovation go with each other; the present exercise raises the question whether public spending in education that seems to have strong effects on urban concentration stimulates such processes. Even more when urban concentration becomes the engine for growth as noted by Kourtit, Nijkamp (2013).

Furthermore, the analysis of data in point 2 has connections with the contributions of Loannides, Overman (2003), Nitsch (2005), Newman (2005), Giesen et al. (2010), Gómez-Déniz et al. (2014), Gómez-Déniz, Calderín-Ojeda (2015), Shujuan (2016), Luckstead, Devadoss (2017), that tried to find the best distributions to fit better the frequency of cities by size. Actually, although deserving further attention in future works, point 2 shows that Pareto distributions seem to fit most countries of the sample, Log normal distributions are better for countries with remote and small cities and a combination of distributions are the outcome of situations when first cities empty second ones.

Finally, this study confirms that truncated data may lead to strange results as it is demonstrated by Jiang et al. (2015) and Cristelli et al. (2012), as for countries with less number of recorded cities, the Zipf elasticity and Intercept do not seem to be interrelated (see Figure 3).

5 Conclusion

The motivation of this paper is to understand what are the effects in urban concentration of the spatial distribution of rents from natural resources and other unilateral transferences. The methodology used assessed the relation between the Zipf curve elasticity by country and indicators of unilateral transferences by country such as the percentage of the rents from natural resources on the national product, the percentage of education and health expenditures on the national product, the percentages of the net FDI outflows on the national product and the percentage of workers' remittances on the national product.

Results show that the ownership over natural resources and the spatial distribution of its rents are decisive factors for urban concentration and regional development. Furthermore, when resource rich countries do not consider this, they miss adopting adequate

policies of spatial allocation of unilateral transferences. Even education policy has a concentration effect. In future work the location of the cities that are more distant from the estimated Zipf curve requires further attention to look into the factors of spatial justice.

References

Ades A, Glaeser E (1995) Trade and circuses: Explaining urban giants. *Quarterly Journal of Economics* 110: 195–228. CrossRef.

Anderson GT, Ying G (2005) The size distribution of Chinese cities. *Regional Science and Urban Economics* 35[2005]: 756–776. CrossRef.

Axtell R, Florida R (2001) Emergent cities: A microeconomic explanation of Zipf's law. Paper presents at the Society for Computational Economics, Yale University

Benguigui L, Blumenfeld-Lieberthal E (2007) Beyond the power law – A new approach to analyze city size distributions. *Computers, Environment and Urban Systems* 31: 648–666. CrossRef.

Bertinelli L, Strobl E (2007) Urbanisation, urban concentration and economic development. *Urban Studies* 44[13]: 2499–2510. CrossRef.

Black D, Henderson V (2003) Urban evolution in the USA. *Journal of Economic Geography* 4: 343–372. CrossRef.

Bosker M, Brakman S, Garretsen H, Schramm M (2008) A century of shocks: The evolution of the German city size distribution 1925-1999. *Regional Science and Urban Economics* 38: 330–347. CrossRef.

Brülhart M, Sbergami F (2009) Agglomeration and growth: Cross-country evidence. *Journal of Urban Economics* 65: 48–63. CrossRef.

Castells M (2012) *Networks of Outrage and Hope. Social Movements in the Internet Age.* Polity Press, Cambridge, MA

Cordoba J (2008) On the distribution of city sizes. *Journal of Urban Economics* 63: 177–197

Córdoba JC (2003) On the distribution of city sizes. working paper, Rice University

Cristelli M, Batty M, Pietronero L (2012) There is more than a Power law in Zipf. *Scientific Reports* 2[1]. CrossRef.

Dentinho TP (2012) New challenges for sustainable growth. In: Capello R, Dentinho TP (eds), *Networks, Space and Competitiveness. Evolving Challenges for Sustainable Growth.* Edward Elgar, Cheltenham, 276–290. CrossRef.

Dentinho TP (2017) Regional development and migration in the lower basin of the Zambezi river. The importance of property rights. *Socio-Economic Planning Sciences* 58: 87–102. CrossRef.

Duranton G (2002) City size distribution as a consequence of the growth process. Department of Geography and Environment, London School of Economics

Eaton J, Eckstein Z (1997) Cities and growth: theory and evidence from France and Japan. *Regional Science and Urban Economics* 27: 443–474. CrossRef.

Gabaix X (1999) Zipf's law for cities: An explanation. *The Quarterly Journal of Economics* 114[3]: 739–767. CrossRef.

Gibrat R (1931) Les inégalités economiques. Librairie du Receuil Sirey, Paris, France

Giesen K, Suedekum J (2014) City age and city size. *European Economic Review* 71: 93–208. CrossRef.

Giesen K, Zimmermann A, Suedekum J (2010) The size distribution across all cities – Double Pareto lognormal strikes. *Journal of Urban Economics* 68: 129–137. CrossRef.

Gómez-Déniz E, Calderín-Ojeda E (2015) On the use of the Pareto ArcTan distribution for describing city size in Australia and New Zealand. *Physica A: Statistical Mechanics and its Applications* 436: 821–832. CrossRef.

Gómez-Déniz E, Sordo MA, Calderín-Ojeda E (2014) The Log-Lindley distribution as an alternative to the beta regression model with applications in insurance. *Insurance: Mathematics and Economics* 54: 49–57. CrossRef.

Henderson JV (2003) The urbanization process and economic growth: The so-what question. *Journal of Economic Growth* 8: 47–71. CrossRef.

Jiang B, Yin J, Liu Q (2015) Zipf's law for all the natural cities around the world. *International Journal of Geographical Information Science* 29[3]: 498–522. CrossRef.

Kourtit K, Nijkamp P (2013) The "New Urban World" – the challenge of cities in decline. *Romanian Journal of Regional Science (Special Issue on the New Urban World)* 7: 10–28

Krugman P (1996) Confronting the mystery of urban hierarchy. *Journal of the Japanese and the International Economies* 10: 399–418. CrossRef.

Loannides YM, Overman HG (2003) Zipf's law for cities: an empirical examination. *Regional Science and Urban Economics* 33: 127–137. CrossRef.

Luckstead J, Devadoss S (2017) Pareto tails and lognormal body of US cities size distribution. *Physica A: Statistical Mechanics and its Applications* 465[Supplement C]: 573–578. CrossRef.

Morudu HD (2016) Developing annual population and gross domestic product estimates for local municipality development models in South Africa: Applications of Zipf's rule from 2001 to 2013. *South African Geographical Journal* 98[2]: 367–385. CrossRef.

Newman MEJ (2005) Power laws, Pareto distributions and Zipf's law. *Contemporary Physics* 46[5]: 323–351. CrossRef.

Nitsch V (2005) Zipf zipped. *Journal of Urban Economics* 57: 86–100. CrossRef.

Peng G (2010) Zipf's law for Chinese cities: Rolling sample regressions. *Physica A: Statistical Mechanics and its Applications* 389[18]: 3804–3813. CrossRef.

Pinto CMA, Lopes AML, Machado JAT (2012) A review of power laws in real life phenomena. *Commun Nonlinear Sci Numer Simulat* 17: 3558–3578. CrossRef.

Quesnay F (1758) O Quadro Económico. Fundação Calouste Gulbenkian, Lisboa

Shujuan L (2016) Rank-size distributions of Chinese cities: macro and micro patterns. *Chinese Geographical Science* 26[5]: 577–588. CrossRef.

United Nations (2014) World Urbanization Prospects: The 2014 Revision, Highlights. (ST/ESA/SER.A/352) Department of Economic and Social Affairs, Population Division

Von Thünen J (1826) *Der Isolierte Staat in Bezienhung auf Landwirtschaft und Nationalökonomie*. Pergamon Press, Oxford. Translated to English by C.M. Wartenberg as "Von Thünen's Isolated State", 1966

World Bank (2010) World Bank Open Data. https://data.worldbank.org/

Zipf GK (1949) *Human Behavior and the Principle of Least Effort*. Addison-Wesley, Cambridge, MA

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 3, 2017, 87–118
DOI: 10.18335/region.v4i3.161

journal homepage: region.ersa.org

Built environment, creativity, social art: The recovery of public space as engine of human development

Anna Onesti[1]

[1] University of Napoli Federico II, Naples, Italy

Received: 11 September 2016/Accepted: 5 September 2017

Abstract. The paper is a part of a comprehensive research aimed at operationalizing HUL approach and experimenting it in the buffer zone of Pompei, mainly in Torre Annunziata (Italy), and is based on the recognition of art and cultural heritage as tools for "managing the change" of landscape. The proposed thesis is that the recovery of public space, configured by art and culture and shared with local community according to an inclusive approach, contribute to regenerate creativity, reconstructing the relationships between people, communities and landscape. This lays the foundations for a "creative environment" and regenerative, concived as a prerequisite of development. In this process, art is a driver which acts on the creativity of local residents, stimulating their critical thinking, open-mindedness and design capacity, and leading them to accept diversity as an opportunity. Focusing on theories and on the empirical analysis of a best practice, MAAM Museum in Rome, this paper has three main objectives: to produce empirical evidence on the relationship between art, heritage and community relationships; to make transferable and replicable in other contexts, such as Torre Annunziata, the process experienced at MAAM; to develop a methodology able to soliciting, integrating and supporting the regeneration of relationships in the town of Torre Annunziata.

1 Introduction

Population growth and migration, climate change and energy resources, economic crisis and social inequality, cultural globalization and rapid growth of ICTs are global processes affecting both people and places. The "New Urban World" (Nijkamp, Kourtit 2012) calls for a deeper reflection on the strategies to implement, by acting locally, with the aim of safeguarding and promoting human wellbeing and human dimension of development. As UNESCO recommendation pointed out, "Urban growth is transforming the essence of many historic urban areas. Global processes have a deep impact on the values attributed by communities to urban areas and their settings, and on the perception and realities of their inhabitants and users" (UNESCO 2011, art. 17).

Despite the inability of many local policies to face new global challenges, all around the world there is an increasing number of bottom-up experiences in which artistic practices activate the change in degraded or abandoned spaces and urban areas, giving shape to new urban landscapes. Slums, abandoned factories, degraded public spaces, historical centers become a testing ground for new forms of synergies between different actors, artists, foundations, cultural associations, third sector organizations, professionals and citizens. From street art to site-specific installations, from poetry to theater up to "live" works, artistic experience, shared with local communities, become instrument to regenerate both

the system of relations between people, which supports the definition of community, and the process of interaction between people and built environment. New forms of patronage, alternative to the public, are emerging and new institutions, legal bodies, associations and foundations arise in order to manage the process of change. Spread throughout the world, these practices are shaping a 'geography of change' (Fondazione Pistoletto 2003), in which art is an instrument of investigation, understanding and change of reality.

It is not just the empirical evidence which underlines the "social" function of art and cultural heritage. By the Faro Convention (Council of Europe 2005b), the European Community recognized that cultural heritage is the fundamental element that characterizes and holds together a community. Through the notion of heritage community (Article 2 b), Faro Convention demonstrates that "by valuing and wishing to pass on specific aspects of the cultural heritage, in interaction with others, an individual becomes part of a community" (Council of Europe 2005a, art. 2). Consistent with the focus definitively on people and "their constantly evolving values, beliefs, knowledge and traditions" (Council of Europe 2005b, art. 2, paragraph a), cultural heritage assumes the widest possible sense, as it can be intended "as a continuing process, of creating, constructing, using and changing heritage" (Fairclough et al. 2014, p. 11). In this framework, promoting cultural heritage protection is a "central factor in mutually supporting objectives of sustainable development, cultural diversity and contemporary creativity" (Council of Europe 2005b, art. 5, paragraph e).

Supporting the principles introduced by Faro Convention, UNESCO Recommendations on Historic Urban Landscape (UNESCO 2011) interpret conservation as a strategy to achieve a balance between urban growth and life quality on a sustainable basis (art. 3) and consider cultural diversity and creativity as key assets for human, social and economic development (art. 12). The definition of tools to implement HUL approach, prompted by UNESCO (2011, art. 24), becomes crucial not only in order to protect the landscape, but above all to build a new "human" development on a local basis, which is consistent with new Millennium goals (United Nations 2015). Operationalizing HUL approach means, therefore, to contribute to achieve the United Nations Sustainable Development Goals and, in particular, the "urban" one (Goal 11) and to make "inclusive, secure, resilient and sustainable cities" (Fusco Girard 2014).

The paper is part of a comprehensive research aimed at operationalizing the HUL approach and experimenting it in the buffer zone of Pompei, mainly in Torre Annunziata (Italy) (Fusco Girard et al. 2016) and focuses on the contribution of art and cultural heritage for "managing the change" of landscape. The proposed thesis is that the recovery of public space, configured by art and culture and shared with local community according to an inclusive approach, contributes to regenerate creativity, reconstructing the relationships between people, communities and landscape. This lays the foundations for a "creative environment" (Törnqvist 1983, Santagata 2009) and regenerative (Fusco Girard 2014), conceived as a prerequisite for development. In this process, art is a driver which acts on the creativity of local residents, stimulating their critical thinking, open-mindedness and design capacity, and leading them to accept diversity as an opportunity.

Focusing on theories and on the empirical analysis of a best practice, such as MAAM Museum in Rome, this paper has three main objectives: to produce empirical evidence on the relationships between art, heritage and community; to make transferable and replicable in other context, such as Torre Annunziata, the process experienced at MAAM; to develop a methodology able to soliciting, integrating and supporting the regeneration of relationships in the town of Torre Annunziata.

2 The beauty of HUL

HUL approach introduces the "principle of relationality" (Fusco Girard 2013) as a new holistic and integrated vision, which links tradition and modernity, past and present, present and future in a systemic/circular and synergistic perspective of development, centered on human beings. This approach outlines a new cultural perspective for urban initiatives, which enhances relationships and interdependencies between the different aspects and the whole and compares universal values and local identities (D'Auria, Pugliese

2013). The historic urban landscape, "world of men and things", can be interpreted as a complex adaptive system with two interacting subsystems: built environment (Ciribini 1979) and heritage community (Council of Europe 2005b). Communities act on built environment through a sedimented intangible cultural capital, making use of and enhancing local resources. In turn, built environment causes relational impacts on communities, regenerating both social ties and their relationships with environment.

This link between cultural production, daily needs and behaviours of people is very strong in Italy: for describing the intangible cultural capital of the community we just say "cultura materiale" ("material" culture), whereas other countries use the term "intangible", focusing more on non-material dimensions (Cuccia, Santagata 2003). Built environment is the tangible expression par excellence of the communities that have shaped it. It reveals in material forms the system of values, exhibiting the processes that, over time, shaped it (Bouchenaki 2003): to meet their needs and to adapt the performance of built heritage – local communities continuously change the landscape (Fontana 2012, Viola 2012), choosing what to preserve and what to change, in a dynamic equilibrium between past and future (Caterina 2012). At the same time, communities are shaped by the built environment which, through its spatial organization and its relationship with nature and climate, defined their structural attributes, influencing the behavior of both individuals and community. The reciprocal interaction between places and communities has produced long-lasting co-evolutionary dynamics (Magnaghi 2012), which in turn created the extraordinary diversity of Italian landscapes. The "co-evolutionary" relationship between community and built environment and the relationship between place and community are both elective relationships of care, feeding, maintenance and culture-expressive: people belong to a territory as they belong to a culture . There can be no landscape without a cultural perspective, allowing the harmonious coexistence of different elements, also very distant in time (Bonesio 2012). As landscape is the creation of a community as a whole, its conservation is related to the reconstruction of community and it is an essential element of local self-sustainable development (Magnaghi 2010).

2.1 The meaning of beauty

"Beauty" has been defined as the set of attributes which man enjoys without desiring to possess them (Eco 2005), going beyond daily needs to express the making of sense and avoiding the fear of passing (Givone 2012). These attributes of beauty makes it ontologically a "common ownership" good which everyone should use without any exclusive claim (Rodotà 2012).

Beauty is not a mere embellishment of reality, but a structural data, which comes from the diversity and plurality of interconnected elements (Fusco Girard 1989). The beauty of landscape derives from the synthesis of different dimensions. Communicating that some criteria/attributes are satisfied at the highest levels, and each element is interconnected to the other, the landscape transfers a sense of wholeness, of fulfillment. In this sense, beauty becomes the main indicator of "proper functioning" of landscape as a complex ecosystem: in the case of a natural landscape, beauty refers to the perfect functioning of the ecosystem, in the case of a built landscape, it shows the link between place and communities, the harmony between nature and man (Fusco Girard, Nijkamp 2005).

A significant contribution to a definition of complex beauty is owed to the theories of the philosopher Arnold Berleant, who pointed out that the loss of human scale in the cities have an aesthetic character, in a sense not limited to visual aspects. As beauty has a synesthetic character, negations of beauty are not only in the transformations which invasively modify scenes and urban views, but also in traffic flows that invade the cities (aesthetic intrusion), in the presence of excessive sound and color (aesthetic distortion), in poor quality of living (aesthetic deprivation), linked to bad sunlight exposure, in unrestrained lifestyle, in the excess of food and consumption (aesthetic depravity). Adopting Schiller's theory (1794 in Berleant 2004) who places the aesthetic experience at the basis of morality, Berleant sees beauty as a source, a sign and a standard of human value. For changing landscape, it is necessary an aesthetic knowledge, meaning the ability to perceive landscape through all the senses, a kind of re-creation, that implies the need to retrace the creative process generating the element which we enjoy (Berleant

2004). Whereas, landscape acts on humans as a field of forces, which establishes an absolute reciprocity relation, so that the user is an integral part of it.

The influence of beauty in physical regeneration, moral and spiritual landscapes and sites have been explicitly recognized by UNESCO since 1962. Beauty contributes to the cultural and artistic life of people but must be harmonized with the needs of communities, their evolution and the rapid development of technical progress (UNESCO 1962). Where settled communities are able to recognize, protect and "produce" beauty issues rather than individual interests, territory has a greater ability to magnetize economic investment, stimulating the new economic activities. A beautiful landscape then becomes a driving factor of economy (Greffe 2005), when the community perceives its value and activates to preserve it. In a circular dynamics, the beauty of the landscape in turn increases the feeling of belonging, the sense of community (Fusco Girard, Nijkamp 2005), and encourages maintenance processes, which result in the conviction that "it is good to" take care of the place. Putting the beauty of the landscape in the center of the transformation process has a double meaning. On the one hand, it means preserving historical beauty, while maintaining efficiency in dynamic processes that shaped the built environment over time as a complex ecosystem. On the other hand, it means reversing the degradation processes and closing the loop between resources and landscape features, enabling new relationships between place and community.

It becomes necessary to "redesign aesthetic of existence", to move beyond the void of appearances, to rethink the unique space of poiesis, its not transitory, its form, its essence, to approach art, not retaining to the past, but saying modernity opposed to conformism (Trione 1996).

2.2 Measurable relationships of landscape

The interactions between community and built environment can be measured by a system of indicators, which allows not only to make communicable and sharable design and selection criteria, but also to determine causal links between heritage and society, while improving the forecasting ability and choice of everyone involved. It allows us to analyze the results of already tested practices, identifying weaknesses to be improved and strengths to be proposed; thereby it facilitates the construction of new tools to support decision making regarding the implementation of UNESCO (2011) recommendations.

The set of indicators must describe the different relationships which feed the landscape system and can be so classified:

1. actions of people on built environment through intangible cultural capital;
2. relations between physical attributes of built environment;
3. influences of built environment on people;
4. internal relationships of heritage community (social capital);
5. relationships of landscape with external environment.

It is interesting to note that there is a close interaction between every single set of indicators in our framework, as every set is both effect of the previous set and cause of the next one. The process of empowerment activated is a regenerative process of development, capable of activating new circular processes, that in turn can face social and physical degradation.

The first series of indicators describes the way in which community is related to built environment through strategies and actions aimed at caring the built environment and at continuing the creative process of landscape production (Magnaghi 2010). In the existing literature, built environment indicators are mainly insufficient and must be improved (Lynch, Mosbah 2017). Indicators related to the attitude of people towards the built environment are useful in order to prove with facts their place-attachment and to make it measurable.

As pointed out by some researchers (Brown et al. 2003, Eshelman, Evans 2002), maintenance and "home personalization" reveal place attachment, whereas observed

incivilities (vandalism, litters and graffiti, ...) predict lower place attachment. The relationship between people and place can be measured also through the attitudes to pro-environmental behaviour (Sanchez, Lafuente 2010, Dunlap et al. 2000), that are strongly linked to values and identity (Gatersleben et al. 2014). Some indicators related to actions of people on built environment have been highlighted in a study aimed at measuring sustainability progress at local level in the United States (Lynch, Mosbah 2017).

Indicators as care, maintenance, recovery and reuse of existing buildings, care and cleanliness of public space and lack of vandalism (Ipsos MORI 2015, Campos, Oliveira 2016), recycling and saving water, saving energy are effective in evaluating the relationship between people and place without resorting to a psychometric approach (Lynch, Mosbah 2017). This paper suggests how to integrate existing indicators with new ones, that better describe the way in which local community acts on the built environment. We suggest to enrich existing literature with new indicators (see Appendix A) as reuse of built heritage, production of site specific art work, integration of art in public spaces, use of public space for artistic activities, rate of local materials and technologies, rate of bioclimatic design solutions, lack of vandalism.

The second kind of indicators describes physical changes produced by the actions of people on place, understood as changes to the structure of the built environment system. This kind of indicators has been mainly studied in the field of visual indicators (Ipsos MORI 2015, Tveit, Sang 2014, Campos, Oliveira 2016), as they are able to describe how much and why a particular place attracts people and activities. Landscape preferences are basically linked to nine visual concepts: naturalness, stewardship, complexity, imageability, visual scale, historicity, coherence, disturbance, ephemera and security (Tveit, Sang 2014). Although our framework excludes some items that are contained elsewhere, indicators related to built environments already proposed in existing literature, as size and quality of public spaces, length of pedestrian paths, preservation status of old buildings, presence of green and open spaces closed to traffic, harmony in dimensional characters (lack of dimensional misalignments in heights), are useful to describe changes in the relationships with the built environment. Aiming to relate built environment to health, an early study conceived by the City Wellbeing Program in Australia (Paine, Thompson 2016) proposes a framework of indicators related to the quality and attributes of built environment, which are very useful in describing internal relationships of built environment system.

Whereas indicators which describe actions of community on built environment and its changes are quite easy to measure, as they can measured mainly through physical dimensions, it is much more difficult to measure the influence of built environment on community and the changes of its internal relations.

Although it is still unclear whether and under what conditions cultural heritage produces beneficial effects on economic development and social community, many areas of impact have been highlighted (D'Auria, Monti 2013).

Starting from research on social capital conducted by Putnam (Putnam et al. 1993, Better Together 2004) and in parallel by Matarasso (1997), it was possible to dissolve the dilemma of whether heritage is "use or ornament". But nevertheless the link between cultural heritage and sustainable development, while described in detail, is still mainly committed to a framework of indicators mostly perceptual and descriptive, based on surveys, interviews, narrative arguments, which cannot determine the causal links between heritage and social growth (Cicerchia 2015).

Earlier studies (CHCFE 2015, HLF 2015) highlight the ability of heritage (and landscape) to enhance both personal and social development. A significant contribution to the identification of social impacts of heritage comes from some studies on the impacts of art on people and communities (Brown 2006, Bollo 2013, Brown, Novak-Leonard 2007, Carnwath, Brown 2014). This field of studies pointed out that cultural participation produces personal and social impacts, not only during the cultural event or immediately afterwards, but also long after the event. In fact, the changes in our beliefs, skills and attitudes rarely are perceived during cultural participation, as they require a sedimentation time before they manifest themselves. The extended impacts, that manifest themselves through behaviors and concrete actions, need a long time (Carnwath, Brown 2014) and

can culminate in other similar impacts. Just cumulative impacts generate significant results in terms of sense of belonging, mindedness, mental health and well-being.

Aiming at describing the influence of built environment on people, individual impacts of heritage can result in three main areas of impact: learning, skills and personal development; improved physical and mental health; cultural activity and well-being (HLF 2015), which increases the level of satisfaction with their lives. Indicators can be distinguish in three relative subcategories. The first one includes indicators as participation in lifelong learning; attending of upper school; students level of literacy and numeracy; level of ICT competencies; specialization in the high knowledge intensity. The second one includes health indicators as life expectancy and healthy life at birth; age-standardised cancer mortality rate (19-64 years old) and mortality rate for dementia and related illnesses (people aged 65 and over); life expectancy without activity limitations at 65 years of age; stress condition. The last one, related to cultural participation, includes involvement of disadvantaged people; incidence of knowledge workers on employment; social and civic participation; voluntary activities.

The systemic structure of social community makes social impacts of heritage more important than the sum of individual impacts. As it improves personal development, in turn heritage induces the improvement of interpersonal relationships. In literature, social impacts of heritage on community relate to three areas of impact: greater interaction between people, that hence the strengthening of social capital; a deeper sense of collective identity, linked to sense of place; enhanced levels of awareness and understanding between particular groups, with a positive effect on community cohesion (HLF 2015).

A significant attempt to make measurable social impacts of heritage on community has been experienced by Dzialek (2014), which, in studying the link between social capital and economic development in different regions of Poland, has brought a number of indicators derived from statistical studies, to three independent components, "formal bridging social capital", "informal bonding social capital" and "informal bridging social capital", stressing the distinction between bonding social capital and bridging social capital (Putnam 2000).

Recently two additional forms of social capital have been added: linking capital, which describes the ability to connect in a vertical direction, and it is the basis of participation around a shared project of individuals with interests and levels of responsibility also very different from each other (Szreter, Woolcock 2004) and can be considered a vertical bridge between powers and asymmetric means (Prior, Tavano Blessi 2012); bracing social capital (Rydin, Holman 2004), used to describe a combination of bridging and bonding capital, but with more attention to the combination of weak and strong bonds in networks. Bracing social capital can be a valid descriptor of hybrid processes of social innovation (Holman, Rydin 2013).

Of special interest are the studies conducted since 2010 by CNEL and ISTAT to measure "fair and sustainable wellbeing" (ISTAT 2015), which integrates indicators of economic, social and environmental measures. The initiative follows the international debate, stimulated by the Stiglitz-Sen-Fitoussi Commission and the international initiatives of the OECD, to measure the progress of societies beyond the PiL (Cicerchia 2015).

This study, that provides every year new data, proposes many useful indicators, able to describe the impact of built environment on the community and the relationships of landscape as a complex system.

The impacts of heritage on community can be expressed as changes in conditions or internal relations, encouraging dialogue between persons not belonging to the same social circle, stimulating acceptance diversity, openness, helping the understanding of different ideas. This type of phenomenon is described through indicators that measure the relatedness of the community and mainly to the presence and density of local associations of various kinds (sports, cultural or social). Other indicators describe the transformation of bridging social capital in bonding social capital and creative growth of the community through the development of cooperative and synergistic initiatives, linked by "working together", which become the glue of society and the engine of attractiveness and competitiveness of the local economy.

The last set of indicators describe the exchanges of information and people with the external environment of landscape as an open system. The possibilities to exchange

are deeply influenced by the connections to the global network (Cohen 2014), that are described through indicators as intensity of use of internet, Wifi coverage, smartphone penetration. The ability of local system to exchange information is described by indicators as numbers of start-up, research intensity, propensity for patenting, rate of technological innovation in the production system, rate of innovation of the product / service of the national production system. As art can be considered a form of communication, the exchange of artworks and products of cultural and creative industries production can be odd as indicators of exchange with external environment too. The second kind of external relationships, that describes the exchange of people, can be measured through indicators as new residents, tourists and visitors, artists, that focus on the presence of foreign people, and on the number of exchange between local community and tourists or artists, that measure more deeply the effective possibilities of change. Last set of indicators provides a final comparison of landscape attractivity through measuring web exposure, popularity index web, prizes and awards. The whole set of indicators is listed in the Appendix A.

The proposed system of indicators, graphically represented on urban maps, leads to a multicriteria approach (Cerreta et al. 2014) for assessing the spatial diffusion of sustainable and human development (Fusco Girard, De Toro 2007, Cerreta, De Toro 2012).

3 Art for the beauty of HUL

3.1 From personal wellbeing to community empowerment

The observation of the multitude of bottom-up initiatives, through the driving force of art, is producing significant results both as recovery of places and regeneration of relationships. Understanding the process by which the individual relationship between person and art contributes to activating a system of relations with the built environment and between people can lead to outline a strategy for urban regeneration processes. Artistic experience is generally a personal emotional experience, which results in a direct relationship between the artistic input and who receives it; from personal emotional experience art enhances wellbeing, also linked at making sense and satisfacting identity needs. From the personal relationship and well-being, art develops an attitude of respect, care, antithetical to the degradation dynamics that characterize the urban spaces. So, the relationship with artwork becomes an attitude of care towards a heritage which is recognized as a common good. Furthermore, sharing the same experience and the same sense of affection, people pass from feeling extraneous in the city to becoming a member of a community (heritage community) (Council of Europe 2005b) whose members recognize the same landscape as cultural heritage. So, a social relationship between people which share the same artistic experience is produced from individual relation between human and artwork. From the regeneration of relations between community and built environment it is possible to activate new virtuous development processes, linked to symbiosis between people/place and synergies between people. The ability of individuals to develop self-identity and autonomous cultural models is reflected in the production capacity of landscape, linked more and more at its ability to offer intangible components. So, art contributes to local development, as it acts on people, influencing their behaviours and openness and stimulating their learning capacity and attitude to innovation (Sacco et al. 2015).

The reconstruction of relations between individuals, community and place, driven by art, lays the foundation of a "creative environment" (Santagata 2009), in which productive synergies and regenerative are activated. These conditions are the requirements of sustainable local development (De Rita, Bonomi 1998, Ciapetti 2010), which relies on the empowerment of local communities and on their ability to produce cultural heritage, managing the change of landscape. In Italy, the white paper of creativity (Santagata 2009) proposes a local development model based on creativity and culture as engines of growth that respects the "grammar of sustainability", as it seeks not only to pass on to future generations infrastructure and cultural resources, but also to preserve equity and defend cultural diversity (Bertacchini, Santagata 2012).

The recognition of the link between creativity, innovation and local development (UNCTAD 2008) suggests to enhance the creativity of people, which is linked to the

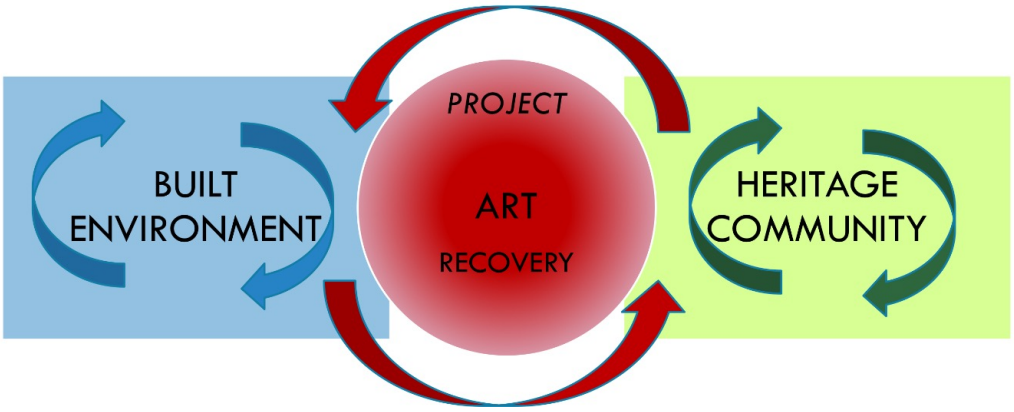


Figure 1: The contribution of art in reactiving the systemic functioning of landscape (Picture: Anna Onesti)

ability to find new and useful combinations between different elements (Poincaré 1906). Creativity of local people is fundamental in building a "creative environment", which is a milieu able to producing and disseminating projects regardless of their scope, either cultural, social, environmental and economic (Greffe 2015). Starting from the consideration that it is possible to distinguish certain activities which play a key role in producing an intrinsic motivation for creative thinking, Sacco, Segre (2009) pointed out the link between creative activities and cultural-led local development, recognizing different "rate" of creativity in human activities. Passing from supercore creativity to creativity, external aims grow up and intrinsic creativity decreases. Art, identified as super-core creativity, has a special function in the relationship between community and built environment, because it contributes to the construction of a new critical knowledge, able to recreate communities. An additional function is performed by the recovery of built heritage, creativity-core activities, which spreads and connects with the built heritage soliciting creative effects from art (see Figure 1).

The art manufacturing process, as well as the recovery of built environment, are both creative activities which, shared with local communities, stimulate individual and collective creativity. While they feed material culture, enhancing planning and designing capacity of local community and putting in relation old place and new technologies, they contribute to regenerating social capital, nurturing both bonding and social capital and building bracing capital. With the transition from personal emotions to community belonging, art contributes to producing new value creation circuits, linked to sharing capacity and cooperative, which in turn introduces new forms of wealth creation. This determines the regeneration of relationships, intended as social capital of community, physical capital of built environment and capacity to change the landscape, in a virtuous regenerative and self-sustaining cycle. As a final outcome, the art determines choral harmony, which generates a field of attractive forces in the landscape. The "beauty" of the landscape, which can be understood as a reflection of an ecology ecosystem between community and built environment, become the main indicator of this harmony. So, it's possible to argue that art contributes to the beauty of landscape not as a decoration but as structural process that guides local communities in the recovery of relationships. Through signs, forms, actions, gestures, the artist invents relationships. Every new artwork is a proposal for how to inhabit a common world (Bourriand 1998).

3.2 A practice: the MAAM, Museum of Other and Elsewhere, in Rome

On March 2009 a group of about 200 people, mainly immigrants, with the support of an organization for the right to housing, occupied an abandoned factory of about 20 thousand square meters in the eastern suburbs of Rome. Few months later, Giorgio De Finis, anthropologist, art curator and filmmaker, with the filmmaker Fabrizio Boni, proposed to the occupants to make a documentario, called "Space Metropoliz", showing

the preparation of a surreal journey to the Moon, the only place available for them. The movie Space Metropoliz (released in 2014) became an opportunity to engage artists and scientists in working with Metropoliziani, teaching them about the moon and building the rocket. At the end of shooting, Metropoliz hosted many artworks, strictly related to the place, and the inhabitants of Metropoliz, feeling "protected" by art, asked to continuing the "game". With more than 500 artworks site-specific and more of 400 artists involved, today they shape Museum of Other and Elsewhere of Metropoliz, one of the most important cultural institutions of Rome (De Finis 2015). MAAM is a new kind of museum, real museum (Pietroiusti 2015) in which people live, taking care of their heritage. Flavours, colours, sounds, art works, historic industrial heritage are integrated in a strange harmony, which confuses art and life in a unique mixed city.

Passing from an "utopia of escape" (to the moon) into an "utopia of reconstruction" (Mumford 1922), MAAM demonstrates that art can be instrumental to build community values, not only interpersonal ones but also interethnic, transforming a disenfranchised group into a heritage community (Council of Europe 2005b). MAAM is recognized by the community as a common good and has established a symbiotic relation between inhabitants and place.

The process through which the art becomes driver of endogenous development can be glances by the following model, that becomes also the reading scheme of the MAAM experience:

- Art as game, Space Metropoliz. The film, presented out of competition at the Venice Film Biennial in 2014, introduces art as both a Trojan to penetrate into the occupied factory and a relational device to liberate the imagination of people and to free them from the burden of everyday life. Intended as a game which makes men free (Schiller 1793 cited in Berleant 2004), art returns to artistic dimension a context without any chance for redemption. Imagining a different world gives Metropoliziani the hope to improve their lives and to build a better future.

- Art as an utopia, a project of change. Passing from Space Metropoliz to MAAM Museum, art guides the transition from an utopia of escape to an utopia of reconstruction (Mumford 1922). Instead of telling the journey of Metropoliziani to the moon as "a happy ending", MAAM shows the physical and social recovery of Metropoliz, which is closer to architects than to artists. Integrating people needs and spacial quality, art contributes to rethink the world and contributes to regenerate social institution and urban community, starting from the recovery of public space, with places for socializing.

- Art as knowledge and communication. Art is instrumental both in learning about the moon and in exhibiting Metropoliz to Romans. As first artworks of MAAM, Gian Maria Tosatti realized a large telescope and placed it on top of the tower, with the collaboration of local inhabitants in cutting, assembling and welding the metal. In a similar way, Hogre made a big sign , with a height of 30 m, which points the way to the moon. As a new urban landmark, these two artworks mark the landscape, referring an idea of art as an ethical sign, which binds matter and thought (Leroy-Gourhan 1964, Lukàcs 1975).

- Art as a barricade. At MAAM, art continuously produces not only cultural values, but also economic, recognized outside, which protect the community from forced eviction and the site from destruction. The demolition of the factory, which would be necessary to the realization of the building complex originally planned, would look more and more like the serious destruction of an artistic valuable collection of hundreds of works recognized by the art system.

- Art as relational system, connecting human activities. At MAAM, art has also another aim: to avoid the enclave effect and open the gates of Metropoliz to the city, sewing together two extremes of the contemporary city, the art museum, the highest place par excellence, and the slum, the lowest and degraded. The attractive power of the art collection of MAAM creates a stream of visitors, who never would have

Table 1: actions on built environment

Maintenance and recovery activities	Diffused in the site
Artworks	500
Production of films	2
Exhibitions	Venere degli stracci by M. Pistoletto held in 2015
Art gallery	1 (Pinacoteca Domestica Diffusa. Diffused exhibition of artworks in houses)
books production	3
Workshop for self-construction	2 (with Facoltà di Architettura Roma Tre 2011; kids rock supported by Commons Camp and Studio Superfluo, 2015)

Figure 2: Local people contributing to the maintenance of artworks

come into contact with this type of community. In this way, art puts in relations Metropoliz to Roma and Tor Sapienza neighborhood, contributing to explain the housing crisis and the serious situation of tens of thousands of people deprived of basic civil rights: school, medical care and voting. Art is in the MAAM a meeting and reporting system to look at others without prejudice.

- Art as a real museum, a new landscape. It develops the idea of a "real" museum (Pietroiusti 2015), an object that exists, has a current consistency in the order of things, concerns facts and existing people, in opposition of the contemporary art museums, unreal objects, seemingly imposed. MAAM can be intended as "place of the Muses", center of knowledge, a place where human activities are connected, as it has been theorized and experimented by Michelangelo Pistoletto (2003).

3.3 Relational indicators at MAAM

The abandonment and degradation that characterized this site before were so strong that social capital had zero value and social dynamics were stationary. In this sense, we can suppose that the social and cultural growth of these communities had no other origin than the process of "recovery" under study.

MAAM can be considered an avant garde for its highly experimental and innovative character, far from ordinary and institutional contexts. For this feature, it can not be replicated elsewhere as it is, but is very useful in order to understanding the relationships that the recovery of public space configured by social art can activate between people, place and community and the processes of involvement of local communities.

Although it was born in an illegal context, free from international policies, MAAM is first of all an approach to local development that, starting from the recovery of public space, acts not only on both physical space but on local people too, as it transforms them in a high quality place and a heritage community strictly linked. MAAM offers a possible answer to some question proposed by United Nations with SDGs as it unconsciously relates to the urban goal, that points out the necessity to make cities more inclusive, secure, sustainable and resilient. MAAM can be considered a good practice for scientific interest because, aiming to rebuild a micro community through the recovery of public space, it tests an innovative tool that although externally-led, as it is a cultural and architectural project, since its birth continuously incorporates community in each phases of development. It is interesting to stress that MAAM came from a specific request of people lived in Metropoliz that, after experimenting the power of art in protecting them

Table 2: Internal relationships of built environment

Urban garden	1
Playroom	1
Schoolroom	1
Sport equipment	2 (football field and basket playground)
Meeting room	1
Piazza	1
Restaurants	1(Cucina Meticcia Metropoliz)

Figure 3: MAAM is part of the skyline of Roma suburbs

Table 3: Internal relationships of heritage community

Local community-culture	200
Local community-people	200
Local community-public	200
Involved disadvantaged people	200
Children involved	70
Afterschool programs	daily
Sport associations	Football team
Sports	Weekly tournment of football; Tournament "Mediterraneo antirazzista"
Educational workshops	1
No profit organizations	2

Figure 4: Art workshop for children in the MAAM ludoteca (playing room)

against forced eviction, proposed to de Finis to continue the "game of art".

Inhabitants of Metropoliz approve every artistic proposals, interact with artists (see Figure 1), sometimes providing material help in building artworks, but never replace artists. MAAM is not an experience of co-design or co-recovery, as it doesn't comes from a local community project, but from a cultural and social project based on community needs and daily shared with them (see Figure 2). MAAM is also supported by third sector organizations and by scholars and cultural associations: in its implementation it proposes a new comprehensive participatory dynamics, which clearly outlines the role and responsibilities of each actor (see Figure 3). Although developed out of the institutional framework, MAAM is an example of hybrid tools for local development through the recovery of public space.

At MAAM art, that is mainly for local people, becomes from a protective barricade a common good, inseparable from the built environment, as people increasingly link their identity to Metropoliz (i.e. they say they are Metropoliziani) and become more and more empowered in place care and management (see Figure 4). Through MAAM museum people of Metropoliz are becoming a heritage community (Council of Europe 2005b), as a social organization that establishes a system of relations and interrelations and adopt a language of its own, through which it communicates with the external environment.

Interactions of people from different countries, ethnicities, social and cultural conditions makes cultural diversity the core elements of MAAM. Although at MAAM art is for

Table 4: Relationships between landscape and external environment

Residents	200
Tourists	Not available
Exchanges tourists/ local community	Open every Saturday
Exchanges artists/local community	Daily
Start-ups	1
Research projects	1
Artworks for export	It is not provided for the production of artworks for sale, just once for self-financing)
Film production	1
New collaborative activities	1 (catering company of local women)
Urban activities	4 (Carnevale a Tor Sapienza; maintenance of urban public space; Free school of Italian for foreigns; La via dell'arte project (artistic path between metro station and MAAM)
Technological innovation for artistic production	3 artworks
Facebook likes	15325
Google	17800 (researching "museo dell'altro e dell'altrove di Metropoliz")
prizes and awards	premio internazionale Marisa Giorgetti. Sezione diritti umani; Institutional recognized in the list of the museums in Rome

Figure 5: The gateway to MAAM

residents, it attracts more and more visitors and it is counted on official Rome website site among main art museums.

Experimenting the proposed framework of relational indicators at MAAM is very useful in order to both understanding and checking the systemic impacts of art on people and place and to testing the application of the framework into a real context.

The first set of indicators (see Table 1), related to the actions of people on built environment, shows very interesting results, as at MAAM more than 500 artworks site specific have been produced (De Finis 2017). Also many cultural events have hosted at MAAM, and are characterized by the strictly integration between cultural activities and public space. This process is accompanied by many maintenance and recovery works and by the reuse of many abandoned space, that are mainly done by local people.

This first set of indicators suggests that art is driver on a new attitude towards built environment.

In turn, this process changes the internal relations of built environment, as it creates new available spaces and new internal paths. The second set of indicators (see Table 2) shows new collective set of spaces: urban garden, playroom, restaurant, meeting room, square and sport equipment.

The third set of indicators (see Table 3) reveals the impact of the process on local people (not only adults but also children and young people) through measuring their active participation to culture and their involvement in the management of Metropoliz. This process of social involvement in turn causes the growth of new collective forms of social capital that, although informal, highlights the production of both bonding and bridging capital. Whereas the first one describes the growth of strong relationships between people, and can be considered the glue of MAAM, the second one refers to weak links between different people, and can be considered the source of MAAM artistic and cultural force. Although the observation of the case study is still too limited in time in

order to demonstrate that the MAAM communities is becoming a heritage community (Council of Europe 2005b), it is noticeable that it is going to build a creative milieu, a social science and cultural context, that is an essential precondition of sustainable development.

The last set of indicators (see Table 4) shows the relationships for MAAM and its external environment through the exchange of knowledge, people and materials. It is interesting to highlight that in the absence of sponsorships and a commercial promotion, MAAM fame was mainly achieved by speaking through the network and unconventional channels. The increasing number of visitors and cultural tourists identifies MAAM as a creative approaches for heritage-based sustainable development (D'Auria 2009). The results achieved at MAAM, expressed in terms of increased relations, synergies and attractiveness, suggest to consider it a best practices. All data refer to 2016.

3.4 Towards an hybrid approach: creative crossovers and "living lab"

The experience of the MAAM can be interpreted as a cultural project, an experiment of social art, which proposes a new dynamic and participatory approach, where the role and responsibilities of everybody are quite clearly outlined. Intended as an example of a cultural project, MAAM paves the way to a public space design aimed at producing cultural crossovers, as systematic methodologically predictable social effects. Cultural crossovers are the effects intentionally produced by the hybridization of art and culture with the most varied sectors, opposite to alternative spillovers, accidental and episodic effects of cultural policy (Sacco, Sciacchitano 2015).

MAAM suggests a process of embeddedness that we propose to implement in different and more ordinary context. In order to make the experience of MAAM replicable and transferable elsewhere, after decoding the process of empowerment of communities through artistic experience, it is necessary to test its "model" in a laboratory context through field testing. Laboratory experimentation allows us to identify cause/effect relationships and to produce empirical evidence about the role of art as a driver of human development.

Live experimentation of the proposed approach requires operational hybrid tools able to meet bottom-up approach of participatory processes with top-down scientific approach and to bring experimentation into a systematic and structured framework of innovation.

Cities are opposite of scientific laboratories that are distinctly and purposefully created to be separate from the lived world in order to manipulating variables and testing hypotheses. Through laboratorization it is possible to set boundaries where controlled experiments can take place and be recorded, in order to transform events/experiments into facts/knowledge. The real world can function as a laboratory, since it adhere to life 'as it is really lived' (Evans, Karvonen 2014). In order to produce laboratory knowledge and make communicable and sharable causal links, it's necessary both to set material, institutional and conceptual boundaries to the testing field and, through measurable indicators, to provide a richness of data that allows statistical patterns to emerge.

An useful tool for hybrid approaches are Living Labs, that were developed in the 90's within MIT, Massachusetts Institute of Technology, with the aim of making knowledge, experiences and daily needs of people the starting point of innovation. Defined by the European network ENoLL (www.www.openlivinglabs.eu) as "user-centred, open innovation ecosystems based on a systematic user co-creation approach integrating research and innovation processes in real life communities and settings", Living Lab is both an approach and an arena (Schliwa 2013) and is characterized by three main features:

1. being confined in a geographically or institutionally bounded space;
2. making social and/or material alterations aimed to conducting intentional experiments;
3. incorporating iterative learning into the process.

Living labs have been exploited as an effective tool in the transitional phases from spontaneous collaboration practices to the empowerment of local communities for local development as they represent a model of territorial innovation based on social economy and community governance (Concilio 2013).

Living lab seems to be a very effective tool that supports the recovery of public space, making it a social innovation and culture-driven tool. As testing arena, Living lab can coincide with public space, that in turn through live experimentation becomes place of social innovation. As place where social processes mainly happen and have a social balancing attitude (Caterina 2013), public spaces become the testing arena for an approach to recovery that actively involves and embeds the whole community.

4 Case study: HUL of Torre Annunziata

The paper focuses on the town of Torre Annunziata, in the Vesuvius area, a case study particularly relevant because it documents a productive landscape (Tempesta 2009), characterized by the symbiosis of a local community dedicated mainly to the production of pasta for food use and a built environment configured for its production (Viola et al. 2014).

Since the nineteenth century, a lot of pasta factories (102 pasta factories; 1.678 employees) prospered thanks to climate attributes (continuous ventilation and good sun exposure), infrastructure (a canal from Sarno river and an industriousness seaport) and the special knowledge of the pasta drying process, imported by a community from Amalfi (Abenante 2011). The urban landscape changed according to the production process; ground floors and basements of residential buildings housed productive uses; a strong link characterized the continuity of public spaces and spatial elements, open up on the street in order to airing and drying pasta and moving raw materials and finished products (Diano 2015). The built environment was configured as a bioclimatic system, which used natural ventilation and rainwater harvesting to facilitate the manufacturing and drying process of pasta production (Napolitano 2015). Airflows, cooled and purified in the transition from green areas, were conveyed into the buildings by means of special architectural concept (Pinto, Viola 2015).

The "quartiere murattiano", Murat district (see Figure 6), developed along two main streets, Via Mazzini and via Oplonti, and bordered by Via Murat and Corso Umberto, hosted the largest concentration of pasta factories, with over one hundred pasta factories in the early twentieth century (Pinto, Viola 2015). Today there is a single pasta factory, the only heir of the local tradition, which preserves a collective memory. The buildings for the production of pasta, arranged along a green area, which favored the ventilation of buildings, have been transformed by various actions that together have altered the character of the area, as well as the green band has been parceled and occupied by a series of poor quality buildings.

The study of Murat district is especially important for knowing dynamics and relationships of the urban landscape system.

The beauty of Torre Annunziata, main indicator of the harmony between communities and built environment, declined with the collapse of pasta production, caused by the interaction of different external pressures. The decommissioning of pasta factories broke the relations between communities and built environment: local community is less and less able to change the built environment and, in turn, built environment produces less and less relational impacts on the community (see Figure 7).

4.1 Relational indicators in Torre Annunziata

In order to understand the systemic functioning of Torre Annunziata landscape, it is useful to measuring relational indicators through the proposed framework. Using the data provided by official sources, integrated with some data derived from on-site analysis, it is possible to have a rather complete picture of the situation (see Appendix B).

The first area of indicators, describing the actions of people on built environment, shows a very critical picture: many are the buildings in mediocre or poor conditions and the index of deterioration is significant, although there is a fairly consistent use of historic buildings.

Place attachment of local community is hard to find: streets and public space are very dirty and there are no public art installations or cultural event, except for the religious

Figure 6: View from buildings roofs in Murat district with Sorrento and its coast in the background

feast Madonna della Neve (Our Lady of Snow), whose procession goes through the center of the town.

The effects of this behaviour on built environment internal relations is obvious, as the second area of indicators shows. In Torre Annunziata there are no public space available for local community and the public green space in the town is very low; incidental contacts between people are blocked more than favored along the streets, that serve only as a park and transit site for cars. All this happens despite the area being subject to legal protection constraints due to the presence of cultural heritage and historic landscape as Oplonti site. These indicators highlight the quality of built environment as a potentiality to improve. Other potentialities are in the bioclimatic functioning of many buildings (ventilation, water recycling system) that, although not in use, can be recovered in order to minimize the consumption of resources.

The influence of built environment on people is very weak: participation in lifelong learning, rate of adults with high school diploma or degree, rate of young people with university education and specialization in the high knowledge intensity are low in Torre Annunziata and lower in Murat district. In turn, the incidence of people not engaged in education, employment or training is quite high. Although they refer to the whole metropolitan area, health indicators as life expectancy at birth and age-standardized cancer and dementia and related illness mortality rate show a low level of well-being. Social participation of people is quite low, as participation rate in election indicators highlight. Volunteers rate and non-profit employees rate are similarly very small, showing low participation of the population in the community social life.

All this results in the scarcity of relationships of heritage community, with few social cooperatives and non-profit organizations.

At the end of our analysis, the last set of indicators, that describes the relationships between landscape and external environment, highlights the whole system criticality. The exchanges of information are very poor as indicators as propensity for patenting, patent impact in innovation sectors reveal. The exportation of art and cultural and creative industries production, that can be considered a form of exchange of information, is virtually zero. The only significant exportation of high quality goods is related to the pasta produced by the only still active pasta factory, that exports its products in USA,

Figure 7: Buit environment and local community decay in Torre Annunziata

UK, France, Spain, Germany, Austria, Greece, Emirates, Japan, Australia.

The exchanges of people, both residents and visitors, are very poor too: demographic variation shows shows a loss of 4490 units from 2001 to 2011 and tourist flow is almost nil, with the exception of visitors to Oplonti archaeological site, which only stay for the time of the visit. Last set of indicators, that provides a final comparison of landscape attractivity by travel reviews on web (TripAdvisor) definitively highlights a very low attractivity of Torre Annunziata and mostly of Murat district.

The whole set of indicators related to Torre Annunziata is listed in Appendix B.

4.2 What of MAAM is replicable in Torre Annunziata?

The recovery of built heritage, while necessary, is not enough to improve the quality of landscape and activate a new development, unless it is not accompanied by the reactivation of local communities and the recovery of their ability to relate to the built environment. This awareness calls for a reflection on cultural and social dynamics capable of contributing to the process of empowerment of local communities.

In order to reactivating the systemic functioning of Torre Annunziata landscape and improving its "beauty", it is necessary to recovery the creativity of its local community, which leads to regenerating the system of relations between people, communities and environment. In order to activating this process, the study suggests to use art as a driver.

We can argue that, first of all, it is replicable the founding idea of MAAM that art, even though cannot bring people on the moon, it can help people to build the moon here on Earth (De Finis 2015), contributing to get people to think differently and to make public space available for the whole community.

It is also replicable the idea that art can be the driver for building a creative milieu, which is the requirement for local development. It's useful to underline that art cannot be considered the characteristic vocation of a place, but the structural element which drives the development and characterizes every place. MAAM is also replicable as a cultural project, in which the physical recovery of public space, configured by art, is instrumentally used to produce social and cultural crossovers: the recovery of public space, as place of relationships, is able to produce social crossovers, depending by both the architectural/artistic design choices and the process of empowerment of local community. MAAM is not a project of self-recovery, but it is a project strongly focused on the need of marginalized communities who occupy the building, which is shared each day.

Though it is impossible to replicate the experience of MAAM, it is possible to reply the process of empowerment, through which art involves local community: it starts from suggesting art as game, an embedding process by which local community can be involved in an utopia, project of the change. This model is based on art as knowledge disposal

and produces a barricade, a system of new values able to protect people and place from external influences, and a relational system, which supports the exchange of information between people. At the end, art makes a new landscape, a real museum, intended as "place of Muses", a place of reproduction of cultural and social values. In its practical implementation, this approach leads to test new forms of landscape management, based on the recognition of public space as a commons.

4.3 Attributes and value of art in three different scenarios

For the development of Torre Annunziata, three different scenarios have been proposed in the PRIN research project (Fusco Girard et al. 2016) developed from Department of Architecture of University of Napoli Federico II: productive networks, touristic hub and centre of sustainable consumption and production. A way to compare them is comparing the meaning of art in each one (see Figure 8).

The first scenario, productive network, sees Torre Annunziata as a place of typical production, strictly linked to tourism. This scenario can include the production of artworks and crafts. The production of art is limited at tourists as customers and art is intended as a private good to take away and is supported by traditional economy, market driven. Art is here instrumental in adding economic values and could produce social spillovers linked to the involvement of local artisans. The second scenario, touristic hub, sees Torre Annunziata as a resort area with accommodation and playground facilities instrumental to increasing residence time of tourists. This scenario can include art both as big event and temporary occurrence and as decoration of touristic places. Art is only for tourists and is intended as a foreign art, imported in order to satisfy the market demand. Art is supported by traditional economy and is a potential source of gentrification. It is instrumental in adding economic values and nurturing local attraction capacity, but it can produce social spillovers, linked to the interaction between residents and tourists, nurturing the sense of identity.

Although the proximity to heritage site and the beauty of Torre Annunziata landscape suggest to focus on cultural heritage as a source of local economy, tourism is not necessary the right way to pursue it, whereas it is one of the fastest-growing economic sectors in the world. Putting tourism at the centre of local development can produce many distortions, as pointed out by an early study (Romão, Nijkamp 2017, WTO 2017). With considering tourism central in local development processes, the recovery of public space must focus mainly on the needs of tourists and only afterwards on the ones of local community. This approach brings to exploit cultural heritage as an economic asset to be exploited in order to meet touristic demand, with a customer oriented strategy mainly based on a oleographic vision that progressively silks its authenticity, cultural vitality and ability to innovate (Sacco et al. 2015). Landscape and cultural heritage are lever for sustainable development, in both economic and environmental, social and cultural terms (Fairclough et al. 2014) as they are vital resources for the citizens (European Commission 2015).

Furthermore, the process of high and quick development of tourism potentially contributes to the reduction of the importance of other economic sectors, including agriculture and manufacturing. In turn, it produces a negative correlation between the rate of workforce employed in tourism and both tourism competitiveness, levels of productivity, resilience (facing economic crisis in 2007), education of population, investments in R&D and attitude to innovation.

But if tourism is programmed and oriented, it can contribute to local sustainable development as it produce a myriad of interactions between insiders, which daily live the place, and outsiders, which see landscape with different eyes. As it contributes to rediscovering the sense of belonging and in turn to re-activating the bond between people and place, this kind of tourism is able to produce circular relationships with local community. The development of tourism related activities must be shared with local community and followed by the development of other economic sectors, able to integrate knowledge, innovation, qualified human resources and value added.

The third scenario, centre of sustainable consumption and production, sees Torre Annunziata as a node in a network of places dedicated to the production and consumption of local goods, as food and crafts. This scenario does not contrast with tourism, but rather

Figure 8: The contribution of art in the proposed scenarios of development

it contributes to develop integrated economic activity and to improve the relationships of tourists and local people. This scenario relates traditional pasta production to the proximity of archeological sites of Villa di Oplonti and Pompeii ruins. Art is intended as daily, permanent opera, produced locally and site-specific. It is mainly for residents though it is a source of attraction for tourists. This kind of art, called social art, is supported by new types of economy as collaborative and sharing economy and is an aid for facing new urban challenges, which are evident in Torre Annunziata. In the third scenario, art is instrumental in adding economic/cultural values and in engaging people, sharing knowledge, connecting people, designing urban recovery, finding the strength for change. It produces social crossovers as it causes the regeneration of relations and inter-relations between people and built environment. The third scenario seems to be the best solution in order to improve the quality of landscape and activate a new development, driven by the reactivation of local communities and the recovery of their ability to relate to natural and built environment.

4.4 Towards a new beauty: art, recovery, public space

In order to implement the chosen scenario, a priority is "regenerating the beauty" of public space, recognized as a space of social capital par excellence (Bullen, Onyx 1998, Prior, Tavano Blessi 2012) in which to activate a new urban metabolism, based on the circularization processes and linked to the symbiotic / synergy concepts (Fusco Girard 2013). This results in actions directed towards the system of urban open spaces and the system of the semi-public spaces, where the exchanges of information, culture, goods, linked to the production and commercialization of pasta take place. The recovery of the public space proposed focuses both on streets, little squares, degraded areas, and entrance hall, courtyards, terraces and, in particular, the green area located between Via Oplonti and Via Mazzini, which today are partly saturated by superfetations and dilapidated buildings. The regeneration of bioclimatic attributes of buildings and urban fabrics, which were once functional to pasta production, can produce added values for both low energy manufacturing and residential and urban comfort. In this sense, the intervention of recovery, which is characterized by the search for a balance between conservation and innovation, can promote a renewed and fruitful dialogue between innovative approaches and local building culture and to activate new synergies and impulses with effects on local economy (Caterina et al. 2015).

In the process of recovering this area, art assumes, as described, a structural function and researches the physical and symbolic integration with architecture. In the past the relationship between art and architecture, called syncretism, was common; in Italy it's still mandatory to allocate a percentage of the budget of public buildings to "make them beauty" through art works (law nr. 717/1945). To this end, the project of recovery must identify spatial and technical elements to make the sublayer and the background of artworks, symbiotically integrated with architecture. The project provides the installation of sculptures and spatial installations along the open area between buildings which, cleared from accretions and shanties, becomes a path in the green, to be used not only as a place

of recreation but also as a slow mobility path. Artworks are also integrated in the street space, configuring the external walls of buildings, especially those with no architectural value, in order to give new meanings to public space and to reconcile differences between new and historical buildings. Although it is mainly focused on the public space, the proposed intervention is completed with the re-use of abandoned buildings by providing, in addition to new production activities, residences and workshops for artists, spread in the urban fabric. The reuse is accompanied by new management mode, providing that, in exchange for their hospitality, the artists are committed to produce and leave some works, whose residents become custodians.

5 Conclusions

The recovery of public space pursues the creativity of local community as a preparatory action to local development and deduces from local community and its sedimented culture the enablers of development.

The recovery process proposed for Torre Annunziata suggests a hybrid approach, in which the "emotional" and community based bottom-up approach with the "rational" and scientific top-down approach are integrated.

In order to operationalise the proposed approach in Torre Annunziata, the paper proposes to use living lab as an hybrid tool able to both experiment in a real field the thesis based on art as driver of development and to really activate new development processes with the empowerment of local communities.

Living Lab is an instrument able to act as an interface between the project and the community in different stages of the process.

Living Lab, implemented in the described area of Torre Annunziata as a real environment, becomes both space of interchange between local knowledge and expert knowledge, way to evaluating relational impacts of the project, place of interaction between institutional actors and new forms of management of common space.

In order to operationalizing the role of art as driver of development, Living Lab assumes the meaning of laboratory of civil aesthetic (Fusco Girard 2012), as a place where to put together artists, designers, citizens, enterprises and associations of the third sector and to produce not only aesthetic values but also social values. The integration between art and built environment, which in the past has always characterized public places, becomes the driver of implementing new forms of communication between culture and communities and regenerating both the material culture, and social capital. Living Lab can be an operational tool for transforming public space into a creative and regenerative environment. In this way, the recovery of public space suggests an alternative/supplementary way to produce wealth and proposes a new idea of economics, as next and cooperative economy, whose core is to participate and to cooperate. It contributes to rebuild a micro-community (or networks of micro-communities), in which economy not only creates but redistributes wealth.

In conclusion, configured by art as driver of development and shared with the local community, the recovery of public space can be proposed as a tool to implement HUL approach, as proposed by UNESCO, and to make cities inclusive, safe, resilient and sustainable. At the same time, the identification of the Living Lab as an auxiliary tool in the implementation of this approach closes the loop with respect to the demand of new operational tools, promoted by UNESCO Recommendations. The approach proposed is also in line with the indications of European Council with regard to participatory governance of cultural heritage (2014/C 463/01) and is consistent with the European Council's conclusion on cultural and creative exchanges to stimulate innovation, economic sustainability and social inclusion (2015/C172/04). In addition, it tracks an operational approach to promote intercultural dialogue and keep the community together through culture in shared public spaces (www.voiceofculture.eu), recognizing that art must be destined to all and not reduced to additional activities for small groups of users.

References

Abenante A (2011) Maccaronari. Ires Campania, Napoli

Berleant A (2004) The aesthetics of art and nature. In: Berleant A, Carlson A (eds), *The Aesthetics of Natural Environments*. Broadview, Peterborough, Ont

Bertacchini E, Santagata W (2012) *Atmosfera creativa. Un modello di sviluppo sostenibile per il Piemonte fondato su cultura e creatività*. Il Mulino, Bologna

Better Together (2004) Bettertogether: The arts and social capital. Saguaro seminar on civic engagement in America, John F. Kennedy School of Government, Harvard University, Cambridge, MA http://www.bettertogether.org/

Bollo A (2013) The learning museum report 3: Measuring museum impacts. http://online.ibc.regione.emilia-romagna.it/I/libri/pdf/LEM3rd-report-measuring-museum-impacts.pdf

Bonesio L (2012) *La questione epistemiologica e il linguaggio. Territorio, luogo, paesaggio*. Firenze University Press, Firenze

Bouchenaki M (2003) The interdependency of the tangible and intangible cultural heritage. 14th ICOMOS General Assembly and International Symposium: 'Place, memory, meaning: preserving intangible values in monuments and sites', Victoria Falls, Zimbabwe. http://openarchive.icomos.org

Bourriand N (1998) *Esthétique relationnelle*. Les presses du réel, Dijon

Brown AS (2006) An architecture of value. *GIA Reader* 17[1]: 18–25

Brown AS, Novak-Leonard J (2007) Assessing the intrinsic impacts of a live performance. Dance Center, Columbia College, Chicago

Brown B, Perkins DD (2001) Neighborhood revitalization & disorder: An intervention evaluation. Final Project Report, National Institute of Justice, Washington, DC, https://www.ncjrs.gov/pdffiles1/nij/grants/196669.pdf

Brown B, Perkins DD, Brown G (2003) Place attachment in a revitalizing neighborhood: Individual and block levels of analysis. *Journal of environmental psychology* 23[3]: 259–271. CrossRef.

Bullen P, Onyx J (1998) Measuring social capital in five communities in nsw: A practitioner's guide. Management alternatives pty limited

Campos A, Oliveira RC (2016) Cluster analysis applied to the evaluation of urban landscape quality. *WIT Transactions on Ecology and the Environment* 204: 93–103

Carnwath JD, Brown A (2014) Understanding the value and impacts of cultural experiences. Arts council England, London

Caterina G (2012) Prefazione. In: Viola S (ed), *Nuove sfide per le città antiche*. Liguori Editore, Napoli

Caterina G (2013) Conservazione, manutenzione e gestione degli spazi pubblici e dei beni architettonici. In: Castagneto F, Fiore V (eds), *Recupero Valorizzazione Manutenzione nei Centri Storici. Un tavolo di confronto interdisciplinare*. Lettera 22, Siracusa, 14–17

Caterina G, Bianchi A, Pinto MR, Viola S, Diano D, Napolitano T, Biancamano PF, Onesti A (2015) A participatory approach for built heritage preservation. Case study: the municipality of Sassano, Italy. In: Amoeda R, Lira S, Pinheiro C (eds), *Proceedings of the 2nd International Conference on Preservation, Maintenance and Rehabilitation of Historical Buildings and Structures*. Porto, Portugal, 463–470

Cerreta M, De Toro P (2012) Urbanization suitability maps: A dynamic spatial decision support system for sustainable land use. *Earth System Dynamics* 3[2]: 157–171. CrossRef.

Cerreta M, De Toro P, Fusco Girard L (2014) Integrated assessment for sustainable choices. *Scienze Regionali* 13[1]: 111–142. CrossRef.

CHCFE (2015) Cultural heritage counts for Europe. CHCFE Consortium, http://www.-encatc.org/culturalheritagecountsforeurope

Ciapetti L (2010) *Lo sviluppo locale.* Il Mulino, Bologna

Cicerchia A (2015) Why we should measure. what we should measure. In: Cicerchia A (ed), *Economia della Cultura.* Rivista dell'Associazione per l'Economia della Cultura, 11–22

Ciribini G (1979) *Introduzione alla tecnologia del design.* Franco Angeli Editore, Milano

Cohen B (2014) Smart city index master indicators survey. Smart cities council, http://-smartcitiescouncil.com/resources/smart-city-index-master-indicators-survey

Concilio G (2013) Innovazione territoriale e living lab. Esperienze e visioni, Cilento Labscape, Working Paper, Napoli

Council of Europe (2005a) Explanatory report to the council of Europe framework convention on the value of cultural heritage for society. Council of Europe treaty series, no. 199, https://rm.coe.int/16800d3814

Council of Europe (2005b) Framework convention on the value of cultural heritage for society (Faro convention). https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/199

Cuccia T, Santagata W (2003) Adhesion - exit: incentivi e diritti di proprietà collettivi nei distretti culturali, in Diritti, Regole, Mercato. Economia pubblica ed analisi economica del diritto. XV Conferenza SIEP, Pavia

D'Auria A (2009) Urban cultural tourism: Creative approaches for heritage-based sustainable development. *International Journal of Sustainable Development* 12: 275–289. CrossRef.

D'Auria A, Monti B (2013) The guardianship of the landscapes between identification and assessment: Ischia and its lost identity. In: Rezekne Higher Education Institut (ed), *Society, integration, education: utopias and dystopias in landscape and cultural mosaic: visions values vulnerability*, Volume IV. Proceedings of the International Scientific Conference on Society, Integration, and Education Location, Udine, Italy, 165–176

D'Auria A, Pugliese S (2013) The governance of UNESCO cultural landscapes between universal values and local identity: the case of Campania. In: Rezekne Higher Education Institut (ed), *Society, integration, education: utopias and dystopias in landscape and cultural mosaic: visions values vulnerability*, Volume V. Proceedings of the International Scientific Conference on Society, Integration, and Education Location, Udine, Italy, 189–200

De Finis G (2015) *Forza tutt*. La barricata dell'arte.* Bordeaux edizioni, Roma

De Finis G (2017) *Maam. Museo dell'Altro e dell'Altrove di Metropoliz˙città meticcia.* Bordeaux edizioni, Roma

De Rita G, Bonomi A (1998) *Manifesto per lo sviluppo locale.* Bollati Boringhieri, Torino

Diano D (2015) Le pressioni perturbative del sistema insediativo di torre annunziata (NA). *BDC – Bollettino Del Centro Calza Bini* 15[1]: 3960. CrossRef.

Dunlap RE, van Liere KD, Mertig AG, Jones RE (2000) New trends in measuring environmental attitudes: Measuring endorsement of the new ecological paradigm: a revisited NEP scale. *Journal of Social Issues* 56: 425–442. CrossRef.

Dzialek J (2014) Is social capital useful for explaining economic development in Polish regions? 962[2]: 177–193. CrossRef.

Eco U (2005) *Storia della bellezza*. Bompiani, Milano

Eshelman PE, Evans GW (2002) Home again: Environmental predictors of place attachment and self-esteem for new retirement community residents. *Journal of Interior Design* 28[1]: 1–9. CrossRef.

European Commission (2015) Getting cultural heritage to work for Europe. Report of the Horizon 2020 expert group on cultural heritage, Directorate General for Research and Innovation, http://bookshop.europa.eu/en/getting-cultural-heritage-to-work-for-europe-pbKI0115128/

Evans J, Karvonen A (2014) 'Give me a laboratory and I will lower your carbon footprint!' – urban laboratories and the governance of low-carbon futures. *International Journal of Urban and Regional Research* 38: 413–430. CrossRef.

Fairclough G, Dragićević-Šešić M, Rogač-Mijatović L, Auclair E, Soini K (2014) The Faro convention, a new paradigm for socially – and culturally – sustainable heritage action? *Culture* 8: 9–19

Fondazione Pistoletto (2003) Geographies of change. A project by Cittadellarte-Fondazione Pistoletto 2003-2017, http://www.geographiesofchange.net/

Fontana C (2012) Re-tracing the systemic approach in architecture, and developing working tools. Methods, Models, Simulations and Approaches Towards a General Theory of Change. Proceedings of the Fifth National Conference of the Italian Systems Society, Marche Polytechnic University, Italy

Fusco Girard L (1989) Qualità e quantità nello sviluppo urbano. i riflessi sul problema della valutazione. In: Fusco Girard L (ed), *Conservazione e sviluppo: la valutazione nella pianificazione fisica*. Franco Angeli, Milano

Fusco Girard L (2012) Creativity and the human sustainable city: Principles and approaches for nurturing city resilience. In: Fusco Girard L, Baycan T (eds), *Sustainable City and Creativity. Promoting Creative Urban Initiatives*. Ashgate, 55–96

Fusco Girard L (2013) Toward a smart sustainable development of port cities/areas: the role of the 'historic urban landscape' approach. *Sustainability* 5[10]: 4329–4348. CrossRef.

Fusco Girard L (2014) Editorial. *BDC – Bollettino Del Centro Calza Bini* 14[2]: 243–249. CrossRef.

Fusco Girard L, Cerreta M, De Toro P, Pinto MR, Viola S (2016) The PRIN research project. the cultural/historic urban landscape: A resource for urban regeneration and local development. Presentation at "The Science of the City", Advanced Brainstorming Carrefour, Napoli

Fusco Girard L, De Toro P (2007) Integrated spatial assessment: A multicriteria approach to sustainable development of cultural and environmental heritage in San Marco dei Cavoti, Italy. *Central European Journal of Operations Research* 15[3]: 281–299. CrossRef.

Fusco Girard L, Nijkamp P (2005) *Energia, bellezza, partecipazione: la sfida della sostenibilità*. Franco Angeli

Fusco Girard L, Torrieri F (2009) Cultural tourism and strategic evaluations: Towards integrated approaches. In: Fusco Girard L, Nijkamp P (eds), *Cultural Tourism and Sustainable Local Development*. Ashgate, London

Gatersleben B, Murtagh N, Abrahamse W (2014) Values, identity and pro-environmental behaviour. *Contemporary Social Science* 9[4]: 374–392. CrossRef.

Givone S (2012) L'origine della creazione artistica. http://www.arte.rai.it/articoli/givone-lorigine-dellispirazione-artistica-3-di-5/22627/default.aspx

Greffe G (2015) Culture and creativity. In: Kakiuchi E, Greffe X (eds), *Culture, Creativity and Cities*. SUIYO-SHA. Available at https://www.researchgate.net/publication/-301603142_Culture_Creativity_and_City

Greffe X (2005) Culture and local development. OECD, Paris

HLF – Heritage Lottery Fund (2015) Values and benefits of heritage: A research review. Heritage lottery fund, https://www.hlf.org.uk/values-and-benefits-heritage

Holman N, Rydin Y (2013) What can social capital tell us about planning under localism? *Local Government Studies* 39[1]: 71–88. CrossRef.

Ipsos MORI (2015) Beauty and social prosperity survey 2015. Respublica London, https://www.ipsos.com/sites/default/files/migrations/en-uk/files/Assets/Docs/-Polls/ipsos-mori-respublica-topline.pdf

ISTAT (2015) Equitable and sustainable well-being. Project to measure equitable and sustainable well-being 2013-2015, http://www.misuredelbenessere.it/

Leroy-Gourhan A (1964) *Le geste et la parole*. Albin Michel, Paris

Lukàcs G (1975) *Cultura e rivoluzione. Saggi 1919-1921*. Newton Compton Editori, Roma

Lynch AJ, Mosbah SM (2017) Improving local measures of sustainability: A study of built-environment indicators in the United States. *Cities* 60: 301–313. CrossRef.

Magnaghi A (2010) *Progetto locale. Verso la coscienza di luogo*. Bollati Boringhieri, Torino

Magnaghi A (2012) Le ragioni di una sfida. In: Magnaghi A (ed), *Il territorio bene comune*. Firenze University Press, Firenze, Italy, 11–30

Matarasso F (1997) Use or ornament? The social impact of participation in the arts, comedia, stroud. http://www.culturenet.cz/res/data/004/000571.pdf

Mumford L (1922) *Storia dell'utopia* (translated and reprinted 2008 ed.). Donzelli editore, Roma

Napolitano T (2015) Sistema edilizio e risorsa idrica. il caso studio di torre annunziata (na). *BDC – Bollettino Del Centro Calza Bini* 15[1]: 61–70. CrossRef.

Nijkamp P, Kourtit K (2012) The 'New Urban Europe': Global challenges and local responses in the urban century. *European Planning Studies* 21[3]: 291–315. CrossRef.

Paine G, Thompson S (2016) Healthy built environment indicators. City Wellbeing Program, CFRC, UNSW, Australia

Pietroiusti C (2015) Possibili caratteristiche di un museo "reale". In: de Finis G, Benincasa F, Facchi A (eds), *Exploit. Come rovesciare il mondo ad arte*. Bordeaux edizioni, Roma, 737–742

Pinto MR, Viola S (2015) Identità sedimentate e nuova prosperità per il paesaggio urbano produttivo. *BDC – Bollettino Del Centro Calza Bini* 15[1]: 71–91. CrossRef.

Pistoletto M (2003) Il terzo paradiso. 2003–2017, http://terzoparadiso.org/

Poincaré JH (1906) *Scienza e metodo* (translated and republished 1997 ed.). Einaudi

Prior J, Tavano Blessi G (2012) Social capital, local communities and culture-led urban regeneration processes: The Sydney Olympic Park experience. *Cosmopolitan Civil Societies Journal* 4[3]: 78–96. CrossRef.

Putnam RD (2000) *Bowling Alone: The Collapse and Revival of American Community.* Simon and Schuster, New York

Putnam RD, Leonardi R, Nanetti R (1993) *Making democracy work: civic traditions in modern Italy.* Princeton University Press, Princeton

Rodotà S (2012) Il valore dei beni comuni. La repubblica, 5 gennaio 2012

Romão J, Nijkamp P (2017) Spatial-economic impacts of tourism on regional development: challenges for Europe (no. 2017_01). Working paper, University of Evora, CEFAGE-UE (Portugal) http://www.cefage.uevora.pt/en/content/download/6410/76704/version/1/-file/2017_01.pdf

Rydin Y, Holman N (2004) Re-evaluating the contribution of social capital in achieving sustainable development. *Local Environment* 9[2]: 117–133. CrossRef.

Sacco PL, Ferilli G, Tavano Blessi G (2015) *Cultura e sviluppo locale: verso il Distretto culturale evoluto.* Il Mulino, Bologna

Sacco PL, Sciacchitano E (2015) Incroci creativi: due conferenze sulla cultura nel semestre di presidenza lettone dell'Unione Europea. Il Giornale delle Fondazioni, http://www.ilgiornaledellefondazioni.com/

Sacco PL, Segre G (2009) Creativity, cultural investment and local development: A new theoretical framework for endogenous growth. In: Fratesi U, Senn L (eds), *Growth and Innovation of Competitive Regions.* Berlin-Heidelberg, Springer-Verlag

Sanchez MJ, Lafuente R (2010) Defining and measuring environmental consciousness. *Revista Internacional de Sociologia* 68[3]: 731–755. CrossRef.

Santagata W (2009) *Libro Bianco sulla Creatività. Per un Modello Italiano di Sviluppo.* Università Bocconi Editore, Milano, Italy

Schliwa GI (2013) Exploring living labs through transition management. challenges and opportunities for sustainable urban transitions. Masterthesis, Lund University, Lund, Sweden

Szreter S, Woolcock M (2004) Health by association? social capital, social theory, and the political economy of public health. *International Journal of Epidemiology* 33[4]: 650–667

Tempesta T (2009) Economia del paesaggio. In: Boggia A, Festa A (eds), *La valutazione del danno ambientale e paesaggistico.* Villa Umbra-Regione Umbria, Perugia

Trione A (1996) *Estetica e Novecento.* Laterza, Bari

Törnqvist G (1983) Creativity and the renewal of regional life. In: Buttimer A (ed), *Creativity and Context: A Seminar Report.* Gleerup, Lund, 91–112

Tveit MS, Sang AO (2014) Landscape assessment in metropolitan areas-developing a visual indicator-based approach. *SPOOL* 1[1]: 301–316

UNCTAD (2008) Creative economy report. http://unctad.org/en/docs/ditc20082cer_en.pdf

UNESCO (1962) Recommendation concerning the safeguarding of beauty and character of landscapes and sites. http://portal.unesco.org/en/ev.php-URL_ID=13067&URL_DO=-DO_TOPIC&URL_SECTION=201.html

UNESCO (2011) Recommendation on the historic urban landscape. UNESCO World Heritage Centre, Resolution 36C/23, Annex, Paris, France

United Nations (2015) Transforming our world: the 2030 agenda for sustainable development. sustainabledevelopment.un.org

Viola S (2012) *Nuove sfide per le città antiche*. Liguori Editore, Napoli

Viola S, Pinto MR, Cecere AM (2014) Recovering ancient settlements: Approaches to negotiation for collective spaces. Proceedings of 40th IAHS World Congress on Housing, Sustainable Housing Construction, Funchal, Madeira, Portugal

WTO – World Tourism Organization (2017) Tourism highlights 2016 edition. Madrid, UNWTO

A Appendix

Table A.1: Relational indicators

indicators	contents	source
actions on built environment		
Cleanes of public space	Washing and sweeping of public space	Campos, Oliveira (2016)
Care of public space	Maintenance, recovery and management of public space (i.e. conservation of materials of sidewalks, decks, flowerbed, street furniture)	Campos, Oliveira (2016)
Care of public space by citizens	Maintenance, recovery and management of public space by citizens	Revised from Campos, Oliveira (2016)
Care of private buildings	Maintenance and recovery of existing buildings	Revised from Campos, Oliveira (2016), Lynch, Mosbah (2017)
Personalization signs	Family names, initials, ornaments on private buildings	Brown, Perkins (2001)
Use of buildings	Rate of buildings occupacy	Revised from Ipsos MORI (2015)
Reuse of historical buildings	Rate of project of building reuse	Revised from Ipsos MORI (2015)
Art production site-specific	Number of artworks locally produced and destined to remain in situ	Revised from Tveit, Sang (2014)
Public art	Integration of art in public space	new
Art and culture in public space	Rate of use of public space for artistic and cultural activities	new
Pro-environmental behaviours - water	Rate of use of recycling and saving water systems	Revised from Lynch, Mosbah (2017)
Pro-environmental behaviours - energy	Rate of use of renewable energy systems	Revised from Lynch, Mosbah (2017)
Bioclimatic design solutions	Rate of use of bioclimatic design solutions	Revised from Lynch, Mosbah (2017)
Use of local materials	Percentage of local materials and technologies	new
Innovation of local and traditional technologies	Rate of innovation of local and traditional technologies	new
Care of the future	Strategic plans for long-term	Fusco Girard, Torrieri (2009)
Relations between physical attributes of built environment		
Walkability Indices	Number of walking routes with a high walkability score / number of all walking routes (%)	Paine, Thompson (2016)
WalkScore	Rating 1-100 describing easy access to places	Paine, Thompson (2016)
Design of building frontages fostering incidental contact	% dwellings fronting streets or walkways that include opportunities for contact with passers-by (eg. seats on a porch, open windows to living areas)	Paine, Thompson (2016)
Design of common areas in buildings fostering incidental contact	% common areas that include spaces with seats and/or to otherwise linger and talk with neighbours	Paine, Thompson (2016)
Public space accessible to the community	% of open space available with unrestricted access	Paine, Thompson (2016), Fusco Girard, Torrieri (2009)
Contact with nature in public space	Area of public open space including vegetation and/or water / area of all public open space (%)	Paine, Thompson (2016)
Contact with nature in public space	Length of streets containing tree plantings: length of all streets (%)	Paine, Thompson (2016)

Continued on next page

Table A.1 – continued from previous page

indicators	contents	source
Consistency of the historic urban fabric	Share inhabited buildings constructed before 1919 and in excellent or good condition than the total of the buildings.	BES_Istat (9.10)
Cultural heritage	Number of archaeological, architectural and museum surveyed in the information system "Risk Map of Cultural Heritage" (MiBAC) per sq km	BES_Istat (9.1)
Public expenditure on cultural heritage	Municipal government spending allocated to functions related to the culture and to the goods per capita.	BES_Istat (9.2)
Influence of built environment on people		
Participation in lifelong learning	People aged 25-64 who participated in education (formal education) and training (non-formal education) / population aged 25-64 * 100	BES_Istat (2.6)
People with at least upper secondary education	Percentage of people aged 25-64 years having completed at least upper secondary education on total people aged 25-64 years.	BES_Istat (2.2)
Level of literacy	Scores obtained in the tests of functional literacy skills of students in the 2nd class of upper secondary education	BES_Istat (2.8)
Level of numeracy	Scores obtained in the tests of numeracy skills of students in the II classes of upper secondary education	BES_Istat (2.9)
People with high level of ICT competencies	Percentage of people aged 16 years and over who can perform at least 5 over the 6 listed operations on the computer on total people aged 16 years and over.	BES_Istat (2.10)
Specialization in the high knowledge intensity	Employed in high-tech manufacturing sectors and those of services to knowledge intensive / total employees * 100.	BES_Istat (11.6)
Life expectancy at birth	Life expectancy expresses the average number of years that a child born in a given calendar year can expect to live if exposed during his whole life to the risks of death observed in the same year at different ages	BES_Istat (1.1)
Healthy life expectancy at birth	Average number of years that a child born in a given calendar year can expect to live in good health on the assumption that the risks of death and perceived health conditions remain constant	BES_Istat (1.2)
Age-standardised cancer mortality rate (19-64 years old	Mortality rate for cancer (initial cause) by five year age groups for people aged 19-64 years, standardized by the Italian 2001 Census population of the same age groups.	BES_Istat (1.7)
Age-standardised mortality rate for dementia and related illnesses (people aged 65 and over)	Mortality rate for nervous system diseases and psychical and behavioural disorders (initial cause) by five year age groups for people aged 65 years and over, standardized by the Italian 2001 Census population of the same age groups.	BES_Istat (1.8)
Life expectancy without activity limitations at 65 years of age	Average number of years that a person aged 65 can expect to live without suffering limitations in daily activities due to health problems, assuming that the risks of death and disability remain constant over time and equal to those observed in a specific calendar year	BES_Istat (1.9)

Continued on next page

Table A.1 – continued from previous page

indicators	contents	source
Cultural participation	Synthetic indicator of the level of cultural participation based on the aggregation of the following indicators: People who have seen videotapes or DVDs in the past 12 months; Percentage of people in the 12 months before the interview have traveled at least once in Cinema, Theatre, Museums and exhibitions, archaeological sites, monuments, concerts (classical music, opera + Other music concerts); Reading: Percentage of people who read the newspaper at least once a week; Percentage of people who have read at least one book in the 12 months preceding the interview; Percentage of people who read a magazine regularly (weekly or periodic)	BES_Istat (2.11)
Involvement of disadvantaged people	Disadvantaged persons involved	new
Incidence of knowledge workers on employment	Employers with college education (ISCED 5-6) in professions Scientific Technology (ISCO 2-3) / Total employed * 100.	BES_Istat (11.3)
Social participation	Based on the aggregation of the following indicators: a) People aged 14 and over who during the last 12 months have participated in meetings of associations (cultural / recreational, ecological, civil rights, for peace); b) People aged 14 and over who during the last 12 months have participated in meetings of trade unions, professional associations or category; c) People aged 14 and over who during the last 12 months have participated in meetings of political parties and / or have worked free for a party d) People aged 14 and over who pay a monthly fee to a club or periodic / sports club e) People aged 14 and over in the last 12 months took part in meetings or initiatives (cultural, sporting, recreational, spiritual) made or promoted by parishes, organizations / religious groups or spiritual	BES_Istat (5.1)
Civic and political participation	Based on the aggregation of the following indicators: 1) People aged 14 and over who talk about politics at least once a week 2) People aged 14 and over who inform policy at least once a week 3) People aged 14 and over who participated in online consultations or voting on social issues (civic) or political (eg. Urban planning, signing a petition) in the last 3 months 4) People aged 14 and over who have read and posted opinions on social or political problems on the web over the past 3 months	BES_Istat (6.2)
Voluntary activities	People aged 14 and over in the last 12 months worked free for associations or voluntary groups / People aged 14 and over * 100	BES_Istat (5.5)
Local community-people	Percentage of people involved in local community activities over the past two years	Dzialek (2014)
Local community-people	Percentage of people involved in public meeting (rather than in the workplace) in the last year	Dzialek (2014)
internal relationships of heritage community		
Local community-foundations	Average number of foundations, associations and public organizations per 10,000 inhabitants	Dzialek (2014)
Local community-art and culture	Average number of members of the arts, music and other cultural groups for 10,000 inhabitants	Dzialek (2014)
Local community-hobby	Average number of members of interest groups and hobby for 10,000 inhabitants	Dzialek (2014)

Continued on next page

Table A.1 – continued from previous page

indicators	contents	source
Local community-sport	Average number of members of sports and recreational groups for 10,000 inhabitants	Dzialek (2014)
NGOs	Number of non-governmental organizations registered for 10,000 inhabitants	Dzialek (2014)
Charitable Organizations	Number of public charities for 10,000 inhabitants	Dzialek (2014)
Religious organizations	Member of religious organizations and church	Dzialek (2014)
Religious organizations	Members of organizations, associations, parties, committees advise, religious groups, unions and other groups to 100 people	Dzialek (2014)
No profit organizations per 10,000 inhabitants	Number of non-profit organizations / total population * 10,000	BES_Istat (5.9)
Social cooperatives every 10,000 inhabitants	Number of social cooperatives / total population * 10,000	BES_Istat (5.10)
Coworking activities	Number of activities and businesses housed in coworking	new
Number of networks	Number of networked production activities	new
Relationships between landscape and external environment		
Intensity of use of the internet	People 16-74 years who used the internet at least once a week over the 12 months preceding the interview / persons of 16-74 years * 100	BES_Istat (11.7)
Numbers of start-up	Number of start-ups / total enterprises	Greffe
Research intensity	Spending on R&S/GDP *100 (Ocse).	BES_Istat (11.1)
Propensity for patenting	Total number of patent applications the European Patent Office (EPO) per million inhabitants	BES_Istat (11.2)
Rate of technological innovation in the production system	Companies that have introduced technological innovations (product and process), organizational and marketing in the three-year period / total number of companies with at least 10 employees * 100	BES_Istat (11.4)
Rate of innovation of the product / service of the national production system	Companies that have introduced product innovations-service over three years / Total Companies with at least 10 employees * 100	BES_Istat (11.5)
Contemporary art production to a foreign market	Number of works of art produced on-siteand destined to the external market	Fusco Girard, Torrieri (2009)
Cultural and creative industries production	Number of cultural and creative industries production	Revised from Fusco Girard, Torrieri (2009)
New residents	number of people who decide to live and develop their own projects in a place (young)	Greffe
Tourists and visitors	Number of tourists and visitors	Greffe
Exchanges tourists / local community	Events involving the interaction between tourists and locals	new
Exchanges artists / community	Events involving the interaction between artists and locals	new
Web exposure	Number of results available on google	new
Popularity index web	number of groups like facebook, approval rating on tripadvisor, foursquare, etc ..	new
Prizes and Awards	Number of awards granted in the last five years	new

B Appendix

Table B.2: Existing relationships in Torre Annunziata

indicators	source	metro-area of Napoli	Napoli municipality	municipality	census area	district	year	positive direction
Actions on built environment								
Residential buildings in mediocre or poor condition	ISTAT Ottomila census				459	124	2011	min
Rate of buildings in poor condition	ISTAT Ottomila census			0,3	0,7		2011	min
Rate of buildings in good conservation status	ISTAT Ottomila census			79.8	59.2		2011	max
Index of building degradation	ISTAT			0.2			2011	min
Consistency of occupied historical homes	ISTAT Ottomila census			22.5	41.1		2011	max
Use of buildings	ISTAT Ottomila census			3.1%	10.1%		2011	min
Cleanes of public space (rate 1-5)	on-site analysis				1		2016	max
Art production site-specific	on-site analysis				0		2016	max
Public art	on-site analysis				0		2016	max
Art and culture in public space	on-site analysis				0		2016	max
Annual municipal electricity production from renewable sources (photovoltaic)	GSE			277.91			2016	max
Annual municipal electricity production from renewable sources (photovoltaic)	ispra			0.59			2014	max
Relations between physical attributes of built environment								
WalkScore	www.-walk-score.-com					78	2017	max
Design of building; frontages fostering incidental contact (%)	on-site analysis					80	2015	max
Public space accessible to the community	on-site analysis					0	2015	max
Public green spaces rate	SIT			0.32			2011	max
Contact with nature in public space (1)	on-site analysis					0	2016	max
Contact with nature in public space (2)	on-site analysis					0	2016	max
Consistency of the historic urban fabric	ISTAT BES	38.2					2011	max
Rate of urban areas with protection constraints	MIBACT					1	2011	max

Continued on next page

Table B.2 – continued from previous page

indicators	source	metro-area of Napoli	Napoli municipality	municipality	census area	district	year	positive direction
Cultural heritage	ISTAT BES			1			2011	max
Influence of built environment on people								
Participation in lifelong learning	ISTAT Ottomila census			4.4			2011	max
Rate of adults with high school diploma or degree	ISTAT Ottomila census			48.6	39.1		2011	max
Rate of young people with university education	ISTAT Ottomila census			18.1	12.2		2011	max
Level of literacy	ISTAT BES		190				2014	max
Level of numeracy	ISTAT BES		191.3				2014	max
Specialization in the high knowledge intensity	ISTAT Ottomila census			32.1	26.4		2011	max
Incidence of people not engaged in education, employment or training (NEET)	ISTAT Ottomila census			24.8	32.7		2011	min
Life expectancy at birth (men)	ISTAT BES		77.6				2013	max
Life expectancy at birth (women)	ISTAT BES		82.4				2013	max
Age-standardised cancer mortality rate (19-64 years old	ISTAT BES		11.1				2013	min
Age-standardised mortality rate for dementia and related illnesses (people aged 65 and over)	ISTAT BES		21.5				2013	min
Participation rate in European elections	ISTAT BES		48.5				2014	max
Participation rate in regional elections	ISTAT BES		61.2				2010	max
Volunteers rate per 100 inhabitants aged 14 and over (%)	ISTAT BES		3				2011	max
Non-profit employees	ISTAT			87			2011	max
Internal relationships of heritage community								
Social cooperatives every 10,000 inhabitants	ISTAT BES		1.3				2011	max
No profit organizations per 10,000 inhabitants	ISTAT BES		18.2				2011	max
Relationships between landscape and external environment								
Propensity for patenting (per 1 million of inhabitants)	ISTAT BES		14.5				2012	max
Patent impact in the High-tech sector (%)	ISTAT BES		23.9				2012	max
Patent impact in the ICT sector (%)	ISTAT BES		24.6				2012	max

Continued on next page

Table B.2 – continued from previous page

indicators	source	metro-area of Napoli	Napoli municipality	municipality	census area	district	year	positive direction
Patent impact in biotechnology sector (%)	ISTAT BES		20.2				2012	max
Contemporary art production to a foreign market	on-site analysis					0	2015	max
High quality production to a foreign market	on-site analysis					1	2015	max
Demographic variation 2011-2001	ISTAT Ottomila census			-4490			2011	min
Travel reviews	Trip Advisor			1097		0	2017	max
Things to do	Trip Advisor			33		0	2017	max

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 3, 2017, 119–131
DOI: 10.18335/region.v4i3.143

journal homepage: region.ersa.org

Research infrastructure, networks of science and regional development – the case of Oskarshamn

Folke Valfrid Snickars[1], Ulf Karlsson[2]

[1] KTH, Stockholm, Sweden
[2] Linköping University, Linköping, Sweden

Received: 15 July 2016/Accepted: 13 September 2017

Abstract. Final disposal of nuclear waste is a global engineering challenge. The Swedish nuclear industry has consequently spent more than thirty years investigating the best sites and technologies for the final storage of nuclear waste. Universities have been involved as experts in this large-scale R&D activity. This has resulted in a well-documented body of knowledge for supporting relevant decision-making. Simultaneously, as a result, global research infrastructure networks have been developed and consequently more than 140 PhD theses have been produced. Eleven of these PhD holders are now full professors.

Based on earlier work on research infrastructures from Lund, Hamburg, and Kiruna, see for instance Snickars, Falck (2015), we have addressed the question of the role of a technical research infrastructure in the development of the fields of engineering and natural science while simultaneously generating regional development. It has provided an opportunity to empirically study the use of research infrastructure in a specialized technology field. At the same time, this study investigates one municipality's efforts to specialize in research without a university in the vicinity.

Do networks of cooperation differ between research groups and research infrastructures? How can a region build its smart specialization on research infrastructure? How can research equipment once belonging to a company be transformed to a public research infrastructure asset?

Our results indicate that research infrastructures such as the ones in Oskarshamn are powerful creators of international research networks. It is possible, although somewhat difficult in view of scattered systems for data provision, to assess their academic and societal impacts. Engineering research has its own networks of university-industry and industry-university interaction where project-based value is cogenerated dynamically. In this study, we have come some way towards empirically analyzing the networks of research cooperation between industry and university using methods of infrastructure and social network analysis.

1 Introduction

Final disposal of nuclear waste is a global engineering challenge with solutions varying by country since the trading of waste is not generally a politically viable solution as it is in other waste management areas. Research infrastructure networks can potentially be used as a basis for the transformation of test sites to complementary uses. Dynamic micro analysis can be performed since there is an ample supply of data, albeit less than

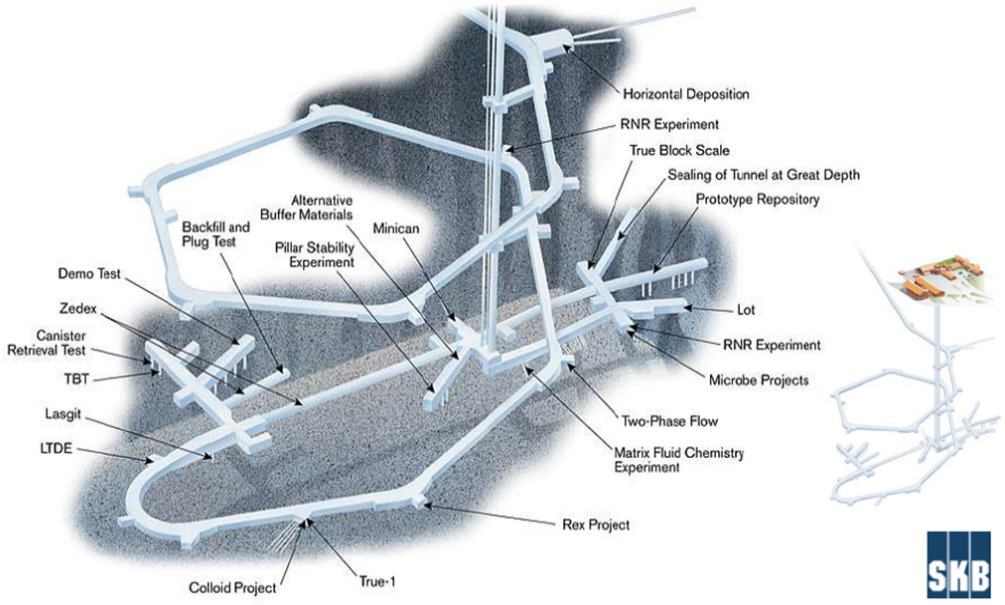


Figure 1: The underground test installations in the Äspö Hard Rock laboratory including names of major experiments

organized. The mutual trading of knowledge between industry and universities involving effective policy schemes can be an interesting spillover mechanism.

Using the Äspö Hard Rock laboratory as a case study, research questions to be addressed in the current paper are as following:

- What quality of science can be generated in an industrial development project?
- How can an experimental research facility belonging to a company be transformed into a public research infrastructure?
- How can a region foster smart specialization through a research infrastructure?

Oskarshamn in southeastern Sweden is the location of one of three nuclear power stations, the primary site for intermediate storage of nuclear waste and the main experimental site for testing technologies for final storage. A decision has been made to store waste in Forsmark northeast of Stockholm for the long-term. National resources are available to assist Oskarshamn in a smooth transformation to another, smarter industrial specialization. The question is, can resources be transformed into new uses in an experimental mine, building on local competence, environmental resources, and international investments in research cooperation?

The core innovation idea is to transform the environment, which is illustrated in Figure 1, to a geosphere laboratory and develop this in the long term as an international research infrastructure in engineering science.

The Äspö laboratory has a depth of some 500 metres and is located on an island within the security zone of the nuclear power plants outside Oskarshamn in a nearby archipelago. Its main mission has been to test the viability and security aspects of storage of nuclear waste. The prototype repository experiment is located in one part of the underground facilities. Other parts are used for the testing of bedrock, and rock drilling properties as well as various environmental experiments to further the understanding of how radioactivity spreads in an underground bedrock environment.

2 The Äspö Hard Rock laboratory in SKB activities

Surveying work began at Äspö in 1986, as the Swedish Nuclear Fuel and Waste Management Company (SKB) wanted to explore the possibility of building an underground laboratory there. This was the beginning of the research and technological development that has taken place in and around Äspö for almost three decades.

It was in its 1986 research, development and demonstration programme that SKB presented the first plans to construct its own underground laboratory. These plans extended from 1987 until 2010 and included for instance detailed studies of natural barriers and their function in a final repository, developing methods and technologies for final disposal as well as demonstrating what its different sections would look like.

The first development measures were already under way before construction of the laboratory had begun. In preparation for, and during construction work, different methods of studying bedrock from the surface were tested. Later the rock was studied in detail in tunnels and shafts. These methods were then used for the comparative site investigations at Forsmark and Oskarshamn before the final choice of Forsmark as the site of the spent fuel repository was made.

The experiments performed at Äspö were collected into a comprehensive database, called Sicada, which now holds the preconditions and results of a 30-year period of experiments, and environmental measurements. The database also comprises information about environmental conditions in the area around the mine itself. It is one of the largest such databases in the world. SKB has also kept track of its research projects which have been organized in a series of three year research and development programmes. There are now around ten programmes in total and documentation on the main results of each programme is currently available.

The activities of SKB can be seen as a large industrial development project. The ultimate goal is to propose a viable and secure technical system to store the nuclear waste in the bedrock forever. A proposal from SKB for this technology is currently being assessed in the Swedish legal system. The assessment will take several years and might lead to a further need for investigations by the Äspö laboratory. The actual depositing of waste is not foreseen to take place before the middle of the 2020s.

The goal of the current paper is to use the publication databases of SKB to estimate some of the scientific and industrial impacts of the research activities at Äspö. The focus will be on publications in a broad sense as they exist in the public databases of SKB. We will characterize the publication patterns reported by SKB in terms of volumes, types of publication, main researchers, and industrial collaboration. In this regard, we will attempt to assess the scientific impact of the work at Äspö for the researchers, and disciplines involved.

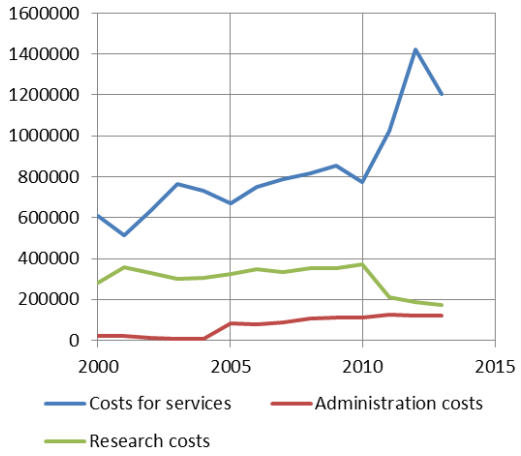
We will provide some basic information about SKB as a starting point for the analysis of publication networks. The presentation will concern SKB as a whole, since it is not possible to separate activities at Äspö from other research activities in the company records. Figure 2 illustrates the total production value of SKB across time.

The total production level was about 1.73 billion SEK in the maximum year 2012. The sharp rise in the production level after 2010 has to do with changes in the company accounting. Even if this is the case the yearly resource use is substantial, for instance, in comparison to the volume of resources used for research in a medium-sized technical university.

Figure 3 illustrates the role of research costs in the company across time. We see a steady decline in the yearly allocations to research both in terms of costs and personnel. The share of research personnel was 40 percent in 2000 and had fallen to 22 percent in 2013. The decline in cost terms is even sharper, reflecting the simple fact that SKB has been approaching the end of its research and development programme.

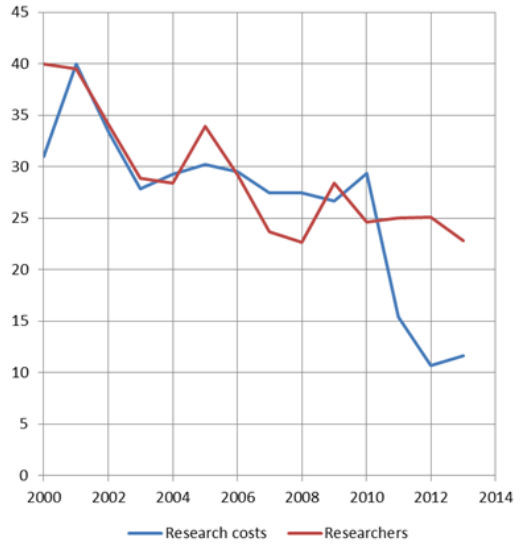
3 Literature review

The analysis is based on earlier work on research infrastructures in Lund, Hamburg, and Kiruna where the emphasis has been on already well-established research infrastructures,



Notes: kSEK current prices

Figure 2: Composition of SKB production system 2000-2013



Notes: in percent

Figure 3: Share of resources allocated to research within SKB 2000-2013)

see Falck et al. (2011) and Snickars, Falck (2015). Those studies have formed part of the input to proposals for new Swedish and EU road maps for research infrastructure raising the question of the importance of those infrastructures for regional development. If such effects can be shown this will add both to the scientific value of investments, as well as the use of EU structural funds and related investment programmes to generate even broader societal impacts.

The Oskarshamn case provides an opportunity to empirically study the use of research infrastructure in a highly specialized technology field where activities are performed in a parallel fashion by university researchers, and development-oriented firms and consultants. In our earlier studies, the starting point has been basic research in physics and the question has been whether this research can be proved to have industrial and societal impacts. In the current study, the starting point is an advanced technological development project in the nuclear field and a central question has been whether this development project has had measurable scientific impacts. A further question is how important and lasting the collaboration has been for the firms, and thus for the impacts on regional development and growth.

The current case study also provides an opportunity to test contemporary theories on the knowledge industry and its mechanisms for the generation and diffusion of knowledge spillovers and innovations. The work is based on thought models in Mellander, Florida (2007), Andersson, Beckmann (2009), Helmers, Overman (2013) and Batabyal, Nijkamp (2015). The case also provides a challenge to study a municipality's efforts to specialize in research and development without a university in the region or even in the vicinity of the region, see also OECD (2012).

Prior work has largely been focused on externalities between firms, or alternatively university-industry linkages, emphasizing localized knowledge spillover effects and functional regions as arenas for knowledge flows, see also Snickars, Falck (2015). Horlings et al. (2012) argue that the existing literature provides no direct empirical evidence demonstrating that impacts actually occur around scientific research facilities and that there is insufficient evidence to support the claim that such investments will, for instance, attract and retain talent and promote innovation.

Helmers, Overman (2013) is one of the first studies to provide empirical evidence on the relationship between agglomeration and a large-scale scientific research infrastructure. Their study considers agglomerative effects of investments in a so-called synchrotron light source in the United Kingdom, and found that the establishment of the Diamond Light Source in Didcot induced clustering of related research activities and increased the research output of nearby organizations, as well as of organizations that did not utilize the facility.

Rekers (2013) has theoretically portrayed the European Spallation Source (ESS) as an enormous addition to the innovation-based economy of the Öresund region. She argues that the anticipated local benefits associated with ESS are tied to the degree of embeddedness of the facility in regional knowledge networks that facilitate localized learning. However, it is also emphasised that innovative work is inherently uncertain, unanticipated and non-linear, where investments do not directly and predictably lead to successful outputs. The counter-argument is rather that impacts can be proactively created.

Johansson, Quigley (2004) argue that networks may provide some or all of the external utility gains derived from agglomeration since knowledge networks that comprise linkages into the global scientific community enlarge the pool of specialized workers that can be considered as interregional sources of new and diverse knowledge.

The research on the impact and utilization of non-large-scale research facilities appears very limited. Falck et al. (2011) have attempted to explicitly demonstrate how the MAX synchrotron facility in Lund is utilized by researchers across the world and whether there is scope to consider MAXlab as a research infrastructure with regional implications. The main result from this study is that it is not possible to consider MAXlab as a regional research facility without considering the importance of its Nordic, European, and global linkages

The current study can also be conceived through the framework of recent attempts to provide a comprehensive map of the pathways to impact of scientific research, see for instance Snickars et al. (2013). Academic impacts have several interconnecting dimensions. A research project, research program, or research infrastructure, can also generate a wide variety of social and economic impacts. These impacts occur in different time scales but also on different spatial scales. Since research is an activity performed in an open global environment, economic and social impacts may also occur regionally, nationally, and internationally. Snickars, Falck (2015) argue that since the Kiruna-based radar station EISCAT has been for a long time identified as a research infrastructure of international importance, its impacts have surely occurred in the global environment. EISCAT has been selected as a top-priority EU research infrastructure.

4 Research and development outputs

SKB details its plans for continued research and technological developments every third year in a special R&D programme which commenced in the 1980's. The most recent one was released in 2013. In this context, however, we will not deal with the scientific substance

Table 1: Äspö publication pattern by type 1993-2013

From	1999	2002	2005	2008	2011	1999
To	2001	2004	2007	2010	2013	2013
Article	13	8	26	4	32	83
Proceedings	22	13	22	1	5	63
Report	113	176	67	52	28	436
Thesis	25	20	23	22	12	102
Total	173	217	138	79	77	684

of this comprehensive research portfolio, but instead concentrate on the networks within science generated by the research, and on the linkages to industrial development activity.

An international review of SKB's research and development efforts was made in 2012, see Nuclear Energy Agency (2012). This review emphasized the importance for SKB to clearly show the chain of events whereby theory is transformed into practice, including how to translate safety analyses into industrial production and application. This work will become increasingly important closer to the start of construction and operation of the final storage facility in Forsmark. The results of this work are currently reviewed and evaluated by the Swedish Radiation Safety Authority (SSM), in a comprehensive licensing process. The NEA review report praises SKB's approach to public outreach work that is described as at the forefront of international practice.

A major part of the research reported in the current paper has been generated by merging publication data from different sources within SKB. The central source of data has been the yearly reports on the publication output from researchers involved in the SKB research programme. These publications consist of both technical reports, proceedings from workshops and conferences, and academic papers published in peer-reviewed journals, see SKB (1994 2014) and SKB (2013).

The composition of these outputs of the research and development projects is summarized in Table 1 which is constructed from yearly reports for the period 1998-2013 and contains some 700 publications. We note that four out of seven of the recorded publications are technical reports and that a rather small number of publications are recorded as articles.

The funding profile of the research done at the Äspö laboratory is shown to decline across time (Figure 2 and 3). This does not necessarily mean that the publication output follows the same pattern. Figure 4 illustrates that the publication level was highest in the period between the end of the 1990s and the middle of the first decade of the 2000s. About 120 publications per year were then produced. The latest ten-year period has exhibited a steady fall in the number of publications as indicated by the relevant yearly reports.

Since research on final disposal is a complex engineering challenge it might be expected that joint publication would be relatively common. Figure 5 illustrates this pattern for the period 1998-2011. We observe that about half of the publications are single-authored with a tendency of this ratio to increase across time. Around one third of the publications have two authors while the share of publications involving three or more authors is one out of five.

We also observe that the share of publications with two authors has remained stable over time. Instead bigger networks of cooperation have been split and more publications have been single-authored. This would seem to indicate that the share of technical reports has also decreased over time as the need for large-scale experimental setups has been reduced.

The total output of PhD degrees has been in the order of seven per year producing a total of 143 dissertations. The PhD degrees have been awarded from a wide set of Swedish and international universities with a focus on the major technical universities in Sweden, especially the Royal Institute of Technology. Four of the dissertations have been produced at non-Swedish universities, see Figure 6.

The dissertations have been an initial part of the career of the persons involved. We

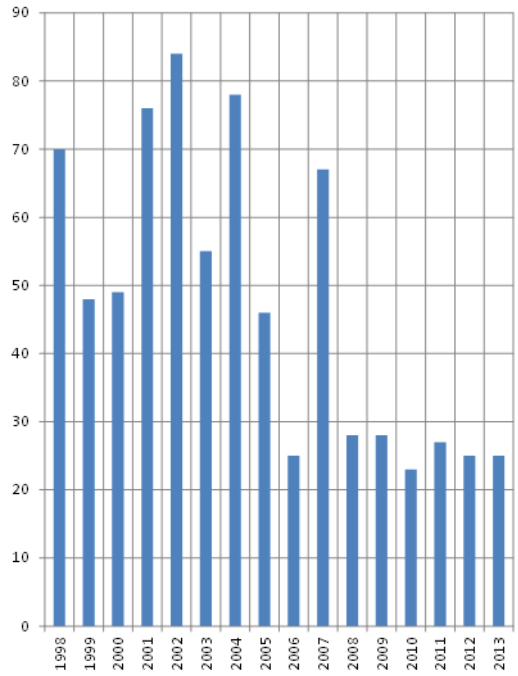


Figure 4: Publications based on experiments at Äspö hard rock laboratory 1998-2013

have followed these careers by tracking down the current affiliation and work title of each of the doctoral students using Äspö as their experimental area, see Figure 7 for an illustration.

In this way, it is possible to characterize an important part of both the academic and the industrial footprint that the Äspö test site has implicitly generated. We will outline the development for dissertations finished during each of the three-year SKB research plans.

We note that eleven of the 143 persons having written their dissertations on results from experimental work performed at the Äspö hard rock laboratory are full professors in 2016. This is somewhat less than one out of ten. Somewhat more than one out of ten has reached director positions in the private or government sector. The bulk of the degree holders have research or management positions. The share of degree holders who are found outside academia increases substantially over time.

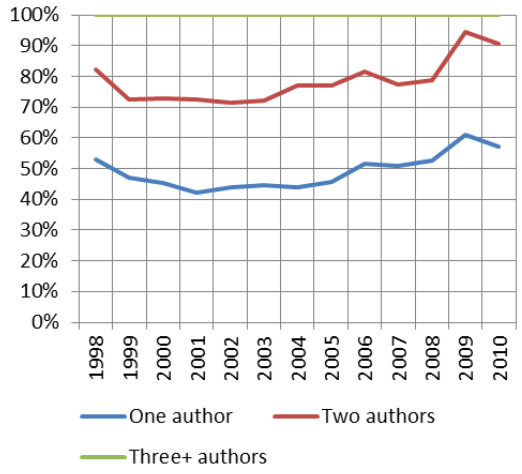


Figure 5: SKB articles according to number of authors across time 1997-2011

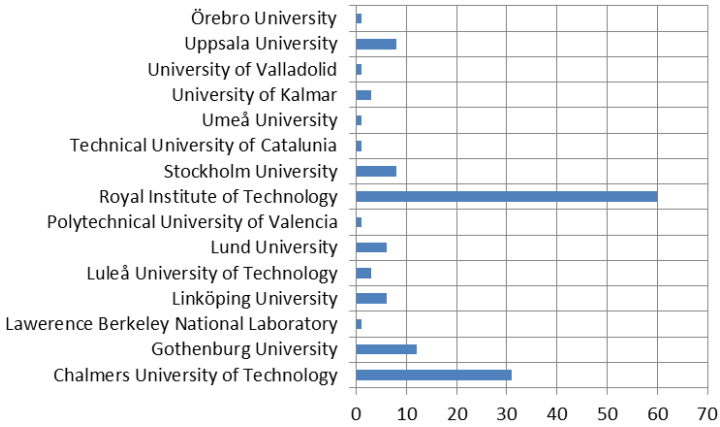


Figure 6: University affiliations of PhDs using Äspö Hard Rock laboratory 1992-2013

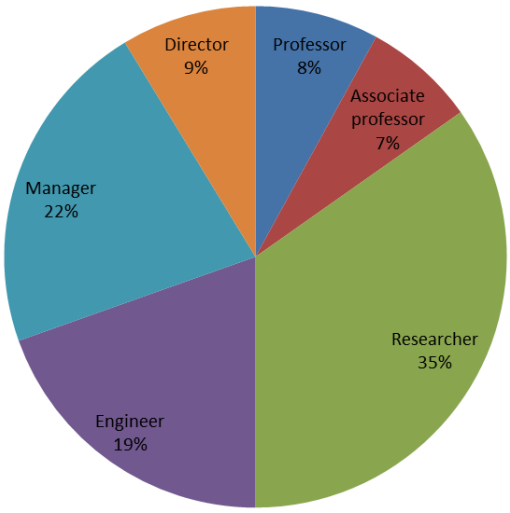


Figure 7: Current work positions for doctoral students using the Äspö hard rock laboratory 1992-2013

One can conclude that an important output of the activities at Äspö has been the production of academic, industrial and management competence for the technology area.

5 Research cooperation and publications

The research and development work at Äspö has been mostly concentrated on testing alternative technologies for the transport, deployment, and storage of nuclear waste. Another focal area has been performing scientific modelling and testing properties of the bedrock in which storage facilities are to be prepared and the waste stored safely for a very long time. For these reasons, documentation of the work has had to be done very carefully so that one can verify results achieved at later points in time.

A comprehensive database, called Sicada, has been set up which also contains systematic measurements of external environmental conditions in the biosphere around Äspö. These arguments indicate that the research and development work has involved a large amount of teamwork, and a large amount of jointly produced publications. However, we have not set up a database for each research project but rather focused on the publications emerging from these projects as they are presented in the yearly reports from SKB.

Figure 8 provides an illustration of patterns of collaboration with non-university organizations as reflected in the publication records from the Äspö hard rock laboratory.

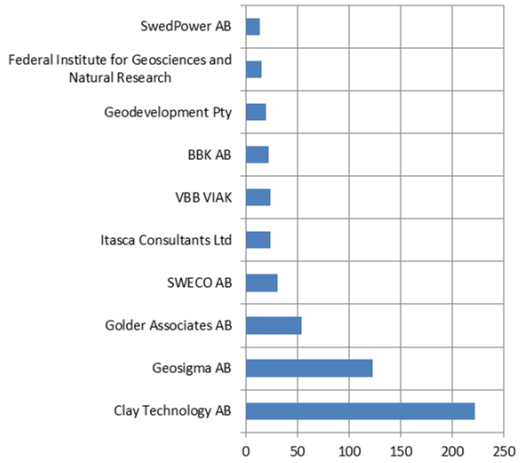


Figure 8: Number of co-authorships involving non-university organizations in publications from Äspö hard rock laboratory 1998-2013

The figure reveals that the most frequently occurring company name is Clay Technologies. It comes as no surprise that this is the company at which two of the leading scientists have been working. The author with the largest number of mentions in the database seems to be working both from his base at Gothenburg University and from his geoengineering consulting company. This observation also holds in a more general sense. It seems that a number of scientists have acted from both university departments and private enterprise platforms in their work at the laboratory.

We can note from the company names shown in Figure 8 that industrial cooperation extends to both Swedish and international companies. Some companies seem to be rather small and very specialized towards the Äspö context. Other networks seem to involve some major Swedish civil engineering consultants who have a specialization in rock mechanics and tunneling.

The pattern observed here is quite different from what may be observed in relation to other research infrastructures. The Äspö case shows that a major technological problem-solving effort can give rise to scientific development, create new academic subjects and provide challenges to existing engineering science.

6 Network analysis of research cooperation

Scientific work depends on the free access to earlier research results. Some publications will be disseminated widely and then effectively function as capital inputs to the ongoing research. The notion of academic impact factors stems from this observation. Highly cited scientists then effectively work as research leaders even if there is no formal connection between the leaders and the followers. In the context of the current case of the Äspö hard rock laboratory we might ask which researchers are the research leaders by analyzing the publication database in terms of patterns of joint authorships. Is it possible to identify subgroups of researchers through the publication analysis?

From the perspective of network theory (Johansson, Quigley 2004), researchers' affiliations are network nodes and co-authorships establishes edges between the nodes. As a result, nodes and edges form pairwise linkages including link-weights for each publication, where linkages indicate joint-publications and link-weights indicates the number of joint-publications between a specific pair. Double back relationships between nodes are controlled for so that each pairwise linkage is unique. The visual exploration of the networks is made using the Gephi software package, see for instance Bastian et al. (2009). The analysis will again be based on publication patterns found in the yearly reports from SKB.

Figure 9 gives an overview of the results. It is obvious that there are groupings of researchers who collaborate with other persons in their fields of science. Some of these

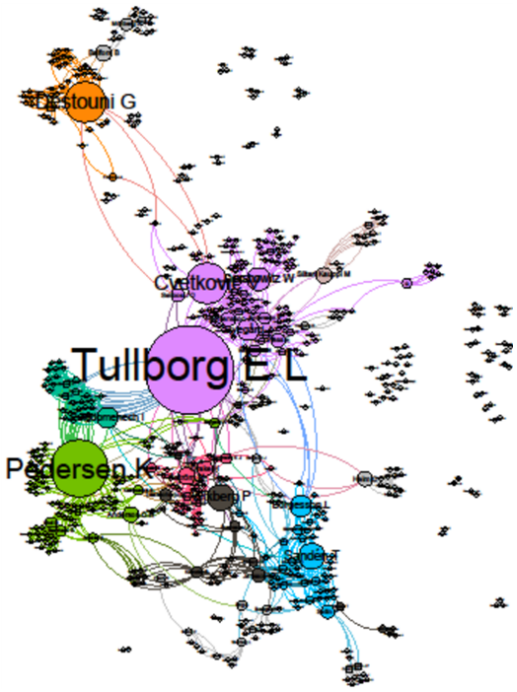


Figure 9: The research network of Äspö-based activities according to SKB yearly reports 1998-2013

groups are independent of the core group, while others are peripheral without belonging to a certain group. In the current case, these are persons who have contributed to the test programmes without having scientific ambitions of their own.

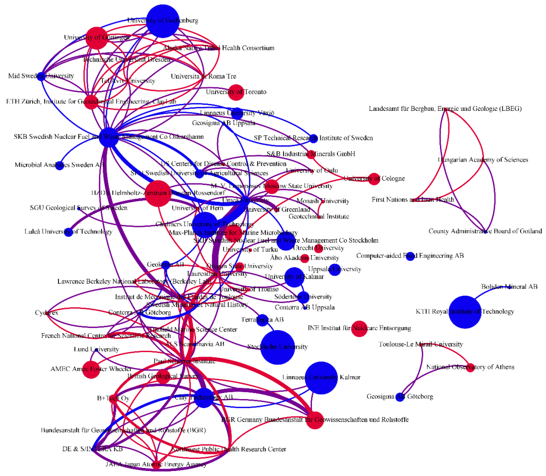
The size and positioning of the persons in the network is determined by the so-called betweenness centrality measure, see Freeman (1977). The indicator measures all the shortest paths between every pair of nodes of the network and then counts how many times a node is on a shortest path between two others. It is a very revealing measure in the case of a network of scientists as it allows us to detect people that occupy an intermediate transfer position between other people or groups.

We note from the figure that some of the researchers who have published the most are not central to the network. Conversely, others stand out as hubs around whom most of the research and development work has emerged during the fifteen-year period investigated. We have partitioned the graph so that those researchers who have the closest affinity in terms of joint publications share the same colour. Seven major groupings have been identified.

Figure 10 shows research networks resulting in academically published articles linked to the Äspö Hard Rock laboratory according to organizational home 1998-2014, see also Falck, Snickars (2017).

As we can see the research activity is characterized by cooperation between researchers from Sweden and other countries. More specifically, among the 90 articles shown in Figure 9 which have been authored by more than 200 persons in more than 70 different research institutions in 18 countries, a small minority do not build on inter-organizational cooperation.

We can conclude that the strategic importance of the Äspö Hard Rock laboratory is not only connected to the task of developing technology for the final disposal of nuclear waste but has also implied strategic research cooperation between researchers in university and industry. This strengthens the position of Sweden as a research nation. The presence of such cooperation accounts for spillover effects in the whole knowledge production system through informal social networks. These effects are not confined to the nuclear industrial networks



Notes: Blue=Swedish organization, Red=International organization. Node size representsnumber of published articlesto which organiations mentioned have contributed. Link thickness indicates number of coauthored and colour (red/blue/purple) 1998-2014

Figure 10: Research networks resulting in academically published articles linked to the Äspö Hard Rock Laboratory according to organization 1998-2014

From a strategic perspective it is not only the presence and volume of research activities which is important but also how the projects have been organized. Projects involving actors from different organizations provide non-monetary benefits which spill-over within the network. The Äspö Hard Rock laboratory has therefore acted as a catalyst for new engineering science and new engineering practices.

7 Conclusions

This study examines the economic and societal importance of the Äspö hard rock laboratory, including the nature, scale, and spatial configuration of related scientific work and other activities. It demonstrates that the nuclear installations in Oskarshamn may give rise to specialized local labour markets, increased human capital levels, consumption of local goods and facilities, and perhaps also functions as a magnet for highly-skilled professionals. It appears, however, that the Äspö laboratory neither resulted in any substantial clustering of related activities nor has it generated any significant economic effects in the short and medium term. Instead, it contributes to strengthen regional competitiveness and attractiveness in a more general sense, suggesting the economic and societal importance of the Äspö laboratory as a knowledge infrastructure is long term and should be interpreted in a strategic context.

There are additional aspects of Oskarshamn as an important node in the nuclear cluster in southeastern Sweden, which includes several complementary organizations such as the encapsulation plant, the central interim storage facility for spent nuclear fuel, the canister laboratory and the bentonite laboratory. The long-term work of SKB in general and the Äspö laboratory in particular has thus implied a large injection of research funding across time in the region.

Another important aspect of the Äspö laboratory is as an interconnecting infrastructure and organization in value creating knowledge networks with global reach. The gains from these network linkages or global knowledge pipelines relate to transactions of formal and informal knowledge and other advantages including memberships in scientific communities that contribute to renewal and dynamism of local knowledge. These network linkages may imply the same type of external utility gains derived from any agglomeration of similar or dissimilar activities, and can be expected to arise from extra-regional sources of new and diverse knowledge.

However, the related development work needs to carefully consider how to promote a high level of absorptive capacity among organizations with related activities. This is in order to maximize the potential benefits of hosting an advanced scientific research infrastructure and to facilitate localized learning and thus contributing to strengthen the nation and region as well as the city of Oskarshamn in a sustainable way.

In terms of the classification of impacts presented in Snickars et al. (2013) a number of observations can be made. First, the main impacts of the Äspö laboratory stay within the technological system related to nuclear technology. It simultaneously creates worldwide academic advancement through the use of innovative equipment, techniques and technologies in a cross-disciplinary setting. The existence of the infrastructure in southeast Sweden has a positive side-effect for the training and teaching of students at academic levels in several Swedish universities engaged in nuclear-related research and education.

The Äspö laboratory as a part of the SKB system has played a role in creating social and economic impacts but there is no distinct profile outside the energy sector. It is clear that SKB's activities increase public engagement and interest in research related to energy in general and nuclear energy in particular.

The fact that this research has been effectively performed in Oskarshamn clearly adds to the role of southeast Sweden as an energy cluster. This aspect of Äspö could be further promoted through deepened studies of the current type for other infrastructures. The geosphere-related development at Äspö may enhance quality of life in view of its attractiveness for cultural tourism.

In summary, it might be said that there is a clear potential to promote social and economic impacts of the energy sector in Oskarshamn especially if one considers the sector as a whole. In this context, it is interesting to observe that the Äspö laboratory can have a very strong role in the national research infrastructure system in Sweden, and at the European level. It can be related to other research infrastructures at the European level for which Sweden has been selected as the host. This international visibility does not seem to fully penetrate to the level of regional and local decision making. There is room for action for instance in relation to the new policy round within the structural funds.

We conclude by providing answers to the research questions posed in the introduction:

- Research infrastructures as the Äspö Hard Rock laboratory are powerful and productive creators of international research and research networks;
- It is possible although technically somewhat difficult to reveal the academic and societal impacts of Äspö-related research and development which become visible if the perspective in broad enough;
- Engineering research has its own peculiar networks of university-industry interaction where value is cogenerated dynamically which makes it a challenge to specialize in it for a peripheral region.

We have come some way in analyzing the productivity of research cooperation using the case of Äspö as a test area but considerably more work needs to be done both to enhance the quality of the background data and to demonstrate the knowledge valorization chains.

Acknowledgement

We acknowledge economic support from the municipality of Oskarshamn in performing the research reported in the current paper. We also thank SKB library staff for allowing access to project and publication data.

References

Andersson ÅE, Beckmann M (2009) *Economics of knowledge: Theory, models and measurements*. Edward Elgar, Cheltenham

Bastian M, Heymann S, Jacomy M (2009) Gephi: An open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media

Batabyal A, Nijkamp P (2015) Creative capital in production, inefficiency, and inequality - a theoretical analysis. Paper presented at an international workshop, Tinbergen Institute, Amsterdam

Falck S, Snickars F (2017) Societal benefits of the Äspö hard rock laboratory. Research report, Department of Urban Planning and Environment, KTH

Falck S, Snickars F, Westlund H (2011) Research hubs in the Baltic sea region: An explorative study about research cooperation and the usage of research facilities in physics and life sciences. Research paper, Royal Institute of Technology (KTH) and the Swedish Research Council (VR)

Freeman L (1977) A set of measures of centrality based on betweenness. *Sociometry* 40: 35–41. CrossRef.

Helmers C, Overman HG (2013) My precious! The location and diffusion of scientific research: Evidence from the synchrotron diamond light source. SERC Discussion Papers, SERCDP0131. Spatial Economics Research Centre (SERC), London School of Economics and Political Science, London

Horlings E, Gurney T, Somers A, van den Besselaar P (2012) The societal footprint of big science – A literature review in support of evidence-based decision making. Working paper 1206, Rathenau Instituut, Den Haag

Johansson B, Quigley JM (2004) Agglomeration and networks in spatial economies. *Papers in Regional Science* 83: 1–13. CrossRef.

Mellander C, Florida R (2007) The creative class or human capital? – Explaining regional development in Sweden. Working paper series in economics and institutions of innovation 79, Centre of Excellence for Science and Innovation Studies, KTH. CrossRef.

Nuclear Energy Agency (2012) The post-closure radiological safety case for a spent-fuel repository in Sweden. An international peer review of the SKB license-application study of March 2011. NEA/RWM/PEER(2012)2, OECD, Paris

OECD (2012) Oecd territorial reviews, småland - blekinge. Research report, OECD, Paris

Rekers JV (2013) The european spallation source (ESS) and the geography of innovation. In: Kaiserfeld T, O'Dell T (eds), *Legitimizing ESS: Big Science as a Collaboration Across Boundaries*. Nordic Academic Press, Lund

SKB (1994–2014) Äspö hard rock laboratory annual report 1993-2014. Technical reports

SKB (2013) RD&D programme 2013. Technical report, TR-13-18

Snickars F, Blaus J, Eriksson T, Reitberger G (2013) Economic and social impacts of kth research: A follow-up study of the research assessment exercise 2012. Research report, VINNOVA and KTH

Snickars F, Falck S (2015) Inter-regional trade in research-based knowledge: The case of the EISCAT radar system. In: Batabyal A, Nijkamp P (eds), *The region and trade: New analytical directions*. World Scientific, London, 227–264. CrossRef.

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 3, 2017, 133–151
DOI: 10.18335/region.v4i3.194

journal homepage: region.ersa.org

Social networks, social satisfaction, and place attachment in the neighborhood

Minou Weijs-Perrée[1], Pauline Van den Berg[1], Theo Arentze[1], Astrid Kemperman[1],

[1] Eindhoven University of Technology, Eindhoven, The Netherlands

Received: 15 May 2017/Accepted: 17 Oct 2017

Abstract. Feeling socially integrated and being satisfied with one's social life are important indicators for happiness and well-being of individuals and for the strength of local communities. The effect of the living environment on social networks and the importance of local social contacts in the neighborhood have been addressed by many studies. However, social satisfaction has received little attention in these studies. The aim of this study is to describe and predict the effect of personal and neighborhood characteristics on social satisfaction mediated by the impact of place attachment and neighborhood networks. A path analysis is used based on survey and diary data collected among 177 respondents between April and May 2014 in Eindhoven and surroundings in the Netherlands. Results show that social characteristics of the neighborhood play an important role in explaining social satisfaction of individuals. In addition, results confirm the importance of participating in social activities and walking or cycling in the neighborhood.

1 Introduction

The satisfaction with the social network and social interactions is an important indicator for the quality of life, health, well-being, and happiness of people (Delmelle et al. 2013, Kawachi, Berkman 2001, Umberson, Montez 2010). Being satisfied with your own social contacts and social network could be described as 'social satisfaction'. Social relations and networks not only provide benefits at the individual level, but also at community level (Scheffert et al. 2008) and even at regional, national, or international level (Siegler 2014). People who are excluded from social relationships and social opportunities are less likely to be satisfied with their social life. Social interactions are important for feeling socially integrated and could decrease feelings of loneliness (Knipscheer et al. 1995). On the other hand, having a larger social network and more social interactions does not necessarily imply that people are more satisfied with their social life (Weijs-Perrée et al. 2015).

Geographical proximity facilitates face-to-face interactions between individuals (Foster et al. 2015). In addition, it is recognized that neighborhood characteristics could influence social network patterns (Cattel 2001). Although social relationships between local residents are usually regarded as weak relations, these relations are very important for local residents (Vermeij 2008). Weak local social relations could contribute to more familiarity, to the attachment to a place, feelings of safety, it may provide a bridge to stronger social relations (Buffel et al. 2011, Vermeij 2008), and probably eventually to higher social satisfaction levels. In addition, it is recognized that people discuss many

important matters with weak ties, because they are knowledgeable or that these weak ties are available at that moment (Small 2013). Although previous studies showed that neighborhood social networks are very important, research into the relation between neighborhood social networks and social satisfaction is still limited.

It is recognized that common memories and feelings about the neighborhood, length of residence, satisfaction with local social contacts, and support between local residents in the neighborhood could increase the attachment to a place (Cramm et al. 2012, Rubinstein, Parmelee 1992, van den Berg et al. 2014). In addition, neighborhood contacts are an important predictor for place attachment (Lewicka 2010). Therefore, if people are more attached to their living environment they will, most likely, be more satisfied with their social life in the neighborhood. Moreover, place attachment is also very important for the well-being and life satisfaction of individuals and it could decrease relocation tendencies (e.g. Greif 2009, Sirgy, Cornwell 2002, Theodori 2001).

It appears from existing literature that physical and social characteristics of the living environment and personal characteristics can influence the social network of individuals (e.g. Maas et al. 2009, van den Berg et al. 2011). However, still little is known about the effect of the living environment on social satisfaction. Knowledge about social satisfaction could give a better understanding to the importance of relationships between local residents, their social network and the impact this has on their attachment to the living environment. This knowledge is relevant to urban planners and policy makers who focus on creating livable and healthy social neighborhoods.

Weijs-Perrée et al. (2015) analyzed the relationships between personal and neighborhood factors, whereby characteristics of the social network, loneliness, and social satisfaction were examined. Their study focused specifically on mobility factors. Characteristics of neighborhood social networks were not included. However, neighborhood contacts are potentially important for feeling socially included, increasing life satisfaction, feeling attached to a place (e.g. Dallago et al. 2009, Livingston et al. 2008), and probably eventually for explaining social satisfaction. Existing knowledge on neighborhood social networks is fragmentary and rarely includes the link with social satisfaction. Therefore, the research objective of this study is to bring all existing concepts together into one comprehensive model and to analyze the (direct and indirect) effects of personal and neighborhood characteristics, place attachment, and social contacts in the neighborhood on social satisfaction.

Data was collected in 2014 in the Eindhoven region in the Netherlands among 177 respondents. Using the data set, a path model was estimated to address this research objective. The remainder of this paper is structured as follows. First, based on a literature review possible relationships are identified between local social contacts, personal and neighborhood characteristics, social satisfaction, and place attachment (Section 2). Next, Section 3 describes the data collection procedure, the sample, and the descriptive statistics. In Section 4, the analysis methods and results are addressed. Finally, Section 5 contains the conclusions and a discussion.

2 Local social network, place attachment and social satisfaction

The social network of an individual can be defined as a network of social relationships with family, friends, and neighbors and the characteristics of these relationships (Croezen 2010). It is recognized that the social network and social interactions are important for the life satisfaction of individuals (Delmelle et al. 2013, Helliwell, Putnam 2004); especially that the quality of, or the satisfaction with these social interactions could increase an individual's well-being (Pinquart, Sörensen 2000). Interacting with neighbors (e.g. exchanging favors and small talks) could increase the well-being (Cramm et al. 2012) and happiness of individuals (Taylor et al. 2001).

The neighborhood is a setting for local social interactions, which are important for the sense of belonging or community and the attachment to the neighborhood (e.g. Lewicka 2010, Vermeij 2008). A higher social cohesion level in the neighborhood can provide greater emotional and instrumental support from neighbors and can lead to more social interactions with neighbors (Windsor et al. 2012). Moreover, residents feel safer in

neighborhoods with a higher social cohesion level than in neighborhoods with lower levels of social cohesion (De Jesus et al. 2010).

Many researchers have studied social networks in the neighborhood and the effects of neighborhood factors on neighborhood contacts. These studies mainly focused on the size and composition of the social network and on the amount of social contact, rather than on the quality of this network (i.e. social satisfaction). For example, Thomése, van Tilburg (2000) and van den Berg et al. (2011) analyzed factors that influence the social network. Their results indicate that social network size and composition are affected by neighborhood characteristics such as the degree of urbanization, age-homogeneity, and the percentage of lower income households.

Social characteristics of a neighborhood, such as the proportion of non-western ethnic minorities and the proportion of people with a low income, have been found to increase the number of social interactions between neighbors (van der Houwen, Kloosterman 2011). A possible explanation is that in more homogeneous neighborhoods, people have more shared common beliefs, values, concerns and interests, and are therefore more likely to interact with each other (e.g. Farrell et al. 2004).

With regard to density, previous studies showed mixed results. According to Delmelle et al. (2013), a denser neighborhood increases face-to-face and spontaneous interactions between local residents. However, this finding is in contrast with findings of Brueckner, Largey (2008), van der Houwen, Kloosterman (2011) and Hanibuchi et al. (2012) who showed that a higher density level negatively affects the number and quality of social interactions in the neighborhood. In less dense areas, people probably have more need to interact with neighbors, because of the low supply of facilities (e.g. cafe's, museums, etc.) in the area (Brueckner, Largey 2008).

Besides neighborhood factors, several effects of personal characteristics on the number of local social interactions were found in previous research such as age, ethnicity, household composition, car ownership and employment (van den Berg et al. 2011, van der Houwen, Kloosterman 2011, Foster et al. 2015). Moreover, people with a higher income have a greater range of resources and therefore probably have more access to social contacts outside of the neighborhood and thus fewer neighborhood-based contacts (Moore et al. 2011). Also, home ownership and length of residence positively influence the knowing of and interacting with neighbors (Guest et al. 2006).

Experiences and memories with the neighborhood and the people who live there, could give a sense of security and familiarity and provide an individual identity. The collection of individual memories and common feelings about a location is also called 'place attachment' (Rubinstein, Parmelee 1992). Previous studies on place attachment showed that characteristics of the living environment affect the attachment to a place (e.g. Greif 2009, Hanibuchi et al. 2012, Livingston et al. 2008, van den Berg et al. 2014). For example, Livingston et al. (2008) argue that social cohesion is the most important neighborhood factor that affects place attachment. If people receive more emotional and instrumental support from neighbors, they probably also feel more attached to a neighborhood. Moreover, Hanibuchi et al. (2012) suggest that people living in urbanized areas feel less attached to their living environment. People who live in a high-density area probably have a lower need to interact with neighbors and are therefore less attached to their neighborhood. In addition, place attachment is also affected by the length of residence in the neighborhood, the satisfaction with the composition of the population, and the satisfaction with social interactions in the neighborhood (van den Berg et al. 2014).

Apart from the living environment, previous research also found influences of personal characteristics on place attachment. Homeowners and people who have a job and a high income were found to feel more attached to their living environment (Brown et al. 2003). These people often have a greater opportunity to live in a place or in a neighborhood where they feel connected or attached with. In addition, the attachment to a neighborhood is related to social networks in the neighborhood (Livingston et al. 2008). Cramm et al. (2012) and Dallago et al. (2009) suggest that a higher social cohesion level in the neighborhood and the quality of local social contacts lead to more emotional and instrumental support among neighbors and a sense of community. The attachment to the

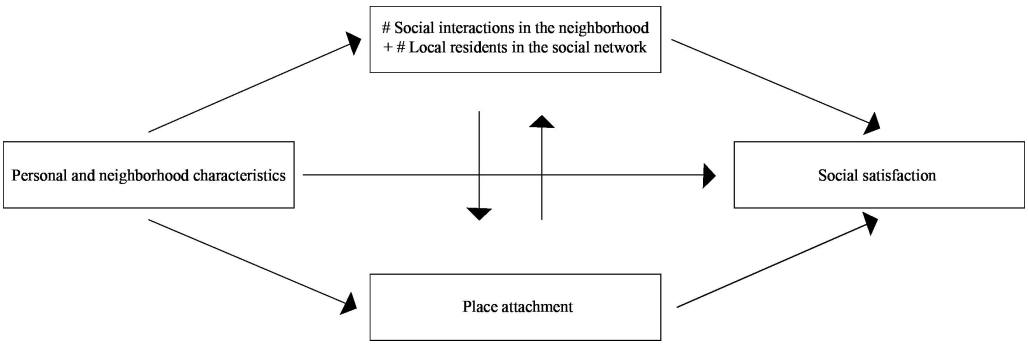


Figure 1: Conceptual model

living environment increases the quality of social relationships with local residents and the sense of community (Dallago et al. 2009). Sense of community is an important factor of quality of life (Kolodinsky et al. 2013). Moreover, if the quality of social interactions increases, the satisfaction with the social life will probably increase.

Most studies on social satisfaction have been conducted by social scientists and have rarely focused on neighborhood characteristics. For example, Bonsang, van Soest (2012) analyzed the determinants of social satisfaction among elderly people in eleven countries in Europe. They found that social satisfaction is affected by income and the participation in non-professional activities. Lansford et al. (1998) suggest that older adults are more satisfied than younger adults with the size of their social network. On the other hand, von Hippel et al. (2008) suggest that aspects of aging, such as spending more time alone, engaging in fewer social activities, and having a poorer working memory, negatively affect social satisfaction.

Social networks and social satisfaction have recently also been studied in travel behavior research (e.g. Carrasco et al. 2008, Delmelle et al. 2013). For example, the frequency of walking in the neighborhood improves health conditions and increases the number of spontaneous and intentional social interactions in the neighborhood (Glanz 2011, van Cauwenberg et al. 2014). Delmelle et al. (2013) found that social satisfaction is affected by personal, neighborhood, and mobility variables (e.g. health status, financial situation, residence time, car ownership, and the urban density). However, these studies did not analyze all concepts into one single comprehensive model.

In summary, the literature review suggests that neighborhood interactions and networks are important indicators for the quality of the social life and the attachment to a place. We hypothesize that, in addition, social satisfaction may also be affected by personal and neighborhood characteristics, mediated by the impact of place attachment and local network characteristics. Although many paths have been reported in previous research, these paths have not been considered simultaneously in a single model or included social satisfaction in relation to neighborhood interactions. Therefore, this study aims to analyze more comprehensively the expected relationships between personal and neighborhood characteristics, the social network and social interactions in the neighborhood, place attachment, and social satisfaction. Figure 1 shows the expected paths, based on the reviewed literature.

3 Data collection, variables and descriptive statistics

To analyze the above-described relationships, a data collection instrument was designed. This data collection instrument consists of a questionnaire on social satisfaction, composition of the social network, place attachment, social cohesion, personal, and neighborhood characteristics and of a social interaction diary to collect data on social interactions. Respondents were asked to fill in the social interaction diary for a limited period of two days to obtain a higher response rate. This data collection instrument is similar to the data collection instrument used by van den Berg et al. (2011). For this study, using a

diary seems to be the most suitable method, because this method has some advantages over other data collection methods. Diaries can be used to collect data on events that are quickly forgotten, about sensitive and personal information, and information about individual's daily behavior or experiences (Corti 1993).

First, the index used to measure social satisfaction is composed of the answers to nine questions on respondents' satisfaction with their social network members (relatives, friends, direct neighbors, local residents, colleagues/ fellow students, club members and other non-kin), social network size and their social life in general. Satisfaction was measured on a 5-point Likert scale ranging from very dissatisfied (1) to very satisfied (5). In the analyses the total score of social satisfaction is used.

Next, the 12-item scale from Williams, Roggenbuck (1989) is used to measure place attachment. This scale contains 12 statements about the connectedness with the neighborhood, namely:

1. I feel that this neighborhood is a part of me
2. This neighborhood is the best place for what I like to do
3. No other neighborhood can compare to this neighborhood
4. This neighborhood is very special to me
5. I identify strongly with this neighborhood
6. I get more satisfaction out of being in this neighborhood than in another neighborhood
7. I am very attached to this neighborhood
8. Doing what I do in this neighborhood is more important to me than doing it in any other place
9. Being in this neighborhood says a lot about who I am
10. I wouldn't substitute any other area for doing the type of things I do in this neighborhood
11. This neighborhood means a lot to me
12. The things I do in this neighborhood I would enjoy doing just as much at a similar neighborhood

The respondents could rate these statements by strongly disagree (1), disagree (2), neutral (3), agree (4) strongly agree (5).

To collect data on social networks a set of name generators was used, which is similar to the name generators used by Carrasco et al. (2008) and van den Berg et al. (2011). The name generators are formulated as:

- Think about the people you feel very close to (people with whom you discuss important matters, keep regularly in touch with or that are there for you if you need help);
- Think about the people you feel somewhat close to (people that are more than just casual acquaintances, but not very close).

This approach was used to identify the number of social network members of respondents in seven social domains: friends, family, direct neighbors, local residents, club members, colleagues/fellow students and other non-kin. In this study, only information about the number of local residents (including the number of direct neighbors) in the social network is relevant.

Finally, information about social interactions was collected using a two-day social interaction diary. Respondents were asked to report social interactions outside of the household, which are more personal than just greetings. Data on social interactions such

as a joint activity (e.g. shopping, dinner or sports), having a conversation (face-to-face, by telephone, through the internet and social media) and sending or receiving a message by email, letter or text message was collected. In addition, respondents were asked to report information about the person they had a social interaction with (e.g. age, gender, social domain, and the strength of the relationship). In this study, only information about the social domain (local residents and direct neighbors) was used to determine the number of social interactions with local residents.

Data on age, gender, household composition, club membership, employment status, health, income, education level, ethnicity, and work status, was collected using a survey. In addition, the frequency of using a car (as driver and as passenger), train, bus/tram/metro, moped or scooter, a bicycle, and walking for transportation was asked. The frequency of using the different transport modes was measured on a seven-point scale ranging from never (1) to almost daily (7).

Regarding neighborhood characteristics, information about density, land use mix, the distance to and the number of facilities, composition of the population, the distance to green public areas, the percentage of non-western/western ethnic minorities and low income groups in the neighborhood, age groups in the neighborhood, and type of dwelling in the neighborhood (stacked or rental dwellings and the average value of per dwelling) were derived from Statistics Netherlands (CBS 2012) using the postal codes of the respondents.

The self-perceived social cohesion was measured using a tool from Frieling (2008). It contains seven questions about social contacts and the degree of solidarity between local residents, namely:

1. How often, in the past six months, did you have a chat with someone from the neighborhood? [once a year or less (1), a few times a year (2), a few times each month (3), once a week (4), a few times each week or more (5)]

2. If you are away from home, is there someone in your neighborhood who looks after your house, for example to make sure that there is no forced entry or give the plants some water? [almost never (1), usually (2), sometimes yes/no (3), usually (4), almost always (5)]

3. If something important happens in the neighborhood or with a neighbor, is there someone in your neighborhood who will make you aware of it? [almost never (1), usually (2), sometimes yes / no (3), usually (4), almost always (5)]

4. Do you feel involved with the people who live in your neighborhood? [with hardly anyone (1), with most people not (2), with some people yes / no (3), with most people though (4), with almost everyone (5)]

5. If there is a sad moment or a sad event in your life, are there local residents who help and support you? [almost never (1), usually (2), sometimes yes / no (3), usually (4), almost always (5)]

6. Are there sometimes any neighborhood parties, barbecues or other activities in the neighborhood, for which the whole neighborhood is invited? [IF YES] How often do you go to these parties, barbecues or activities? [almost never (1), usually (2), sometimes yes / no (3), usually (4), almost always (5)]

7. Have you in the past year collaborated with other local residents to organize something in the neighborhood, for example, to organize a neighborhood party or activity, or to make a neighborhood newspaper [IF YES] How often have you met in the past year with these local residents? [not collaborated (1), collaborated about once every half a year (2), collaborated about once every three months (3), collaborated about once every two months (4), collaborated about every month or more frequently (5)]

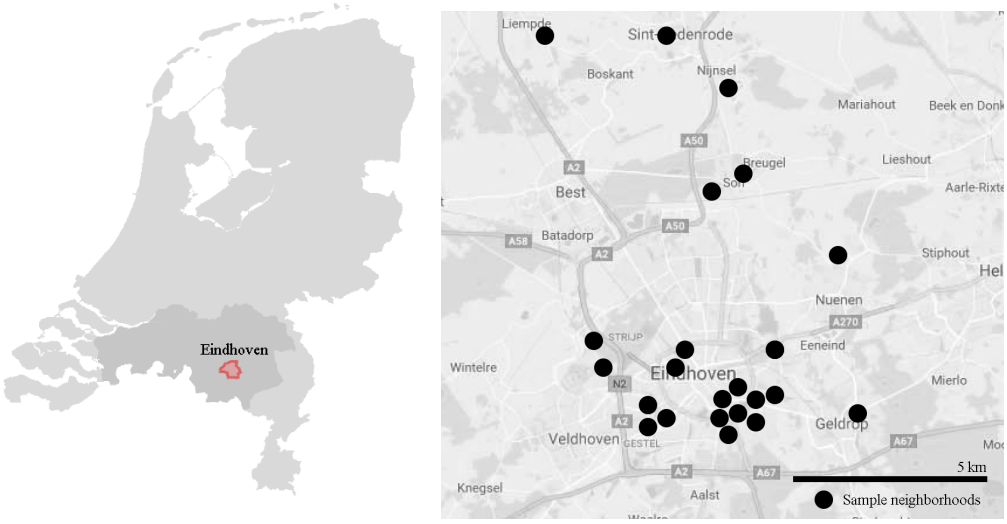


Figure 2: Sampled neighborhoods in Eindhoven and surrounding towns

3.1 Data collection

The aim of this dataset was to collect data among the same respondents that took part in a study of van den Berg et al. (2011) in 2008. In 2008, respondents were randomly selected in the Eindhoven and surrounding towns (Nuenen, Gerwen, Geldrop, Son en Breugel, Liempde, Nijnsel and Sint Oedenrode; see Figure 2). A total of 747 respondents participated in 2008 and 523 addresses of these respondents were known. Between April and May 2014, these 523 respondents were personally approached at their home and the respondents who were not at home, were called subsequently. If people were willing to participate, they received a social interaction diary, which was collected a week later. A total of 141 useful (27% of the 523 respondents that participated in 2008) diaries returned. The overall response rate is only 18% of the randomly selected 747 residents in 2008. Besides these addresses, 47 new respondents were selected, which consist of acquaintances and some randomly selected addresses and a total of 36 useful diaries returned. This resulted in a total of 177 respondents. Out of these respondents, eight respondents only filled in the questionnaire or kept the diary for only one day.

3.2 Descriptive statistics

Table 1 shows the basic sample characteristics of the 177 respondents. As can be seen, the sample contains a higher percentage of women than men. Next, the sample also contains a higher percentage of people aged older than 65 years compared to the Dutch population, municipality of Eindhoven (Gemeente Eindhoven 2017), and the sample of van den Berg et al. (2011). Women and older people probably spend more time at home, so it is more likely that they would answer the door and participate in this study. The sample contains a high percentage of people with a higher education, which is also observed in the sample of 2008. This is probably due to the fact that people with a higher education level are more willing to participate in a research. Compared to the sample in 2008, this sample has a high percentage of people who do not work. These people spend more time at home and therefore probably have more time to participate in this research.

In addition, 18% of the respondents in the sample walk (almost) every day for transportation in the neighborhood and only 9% never walk for transportation in the neighborhood. Almost half of the respondents cycle (almost) every day (42%).

It is assumed that social satisfaction is an interval variable and the total score of the nine items on social satisfaction will be used in the analyses. This total score consists of the score for satisfaction with the different social categories (i.e. friends, family, direct neighbors, local residents, colleagues/fellow students, and club members), satisfaction with the size of the social network, and satisfaction with the social network in general.

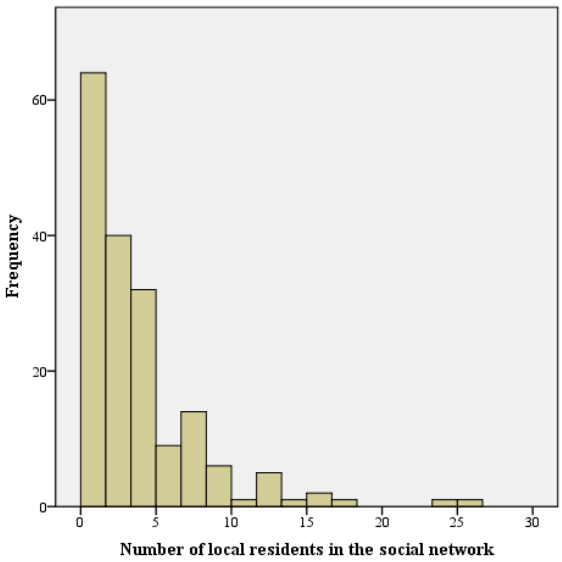


Figure 3: Distribution of the number of local residents in the social network

The Cronbach's Alpha for the total score of social satisfaction is 0.758, which means that the internal consistency of the items is good. The mean total score of social satisfaction is 34.63 and the mean on a five point Likert scale is 3.8.

The sum score of the 12 items for measuring place attachment was used in the analyses. The internal consistency of the items is good; with a Cronbach's Alpha of 0.908. As can been seen in Table 2, the mean total score of the self-rated place attachment is 38.56, with a standard deviation of 7.37. In addition, the average number of local residents (including direct neighbors) in the social network of the respondents is 3.56. Figure 3 shows the distribution of the number of neighbors in the social network. The high number of local residents in the social network is probably due to the fact that the respondents in this study lived in the neighborhood for a longer time. In addition, this could be related to the fact that people who have more social interactions (with neighbors) are probably more willing to participate in this study. This should be taken into account when interpreting the results. The average number of social interactions with local residents per respondent in two days is 0.96, with a standard deviation of 1.48. Figure 4 shows the distribution of the number of interactions with neighbors during the two days of the interaction diary.

The following five different classes of urban density are distinguished by the Dutch Bureau of Statistics (CBS 2012), namely:

1. Very high density (2500 or more addresses per km2)
2. High density (1500 to 2500 addresses per km2)
3. Moderate density (1000 to 1500 addresses per km2)
4. Low density (500 to 1000 addresses per km2)
5. Very low density (less than 500 addresses per km2)

Most of the respondents in the sample live in a neighborhood with more than 1000 addresses per km2 [(35%) moderate density, (21%) high density, or (23%) very high density]. Only 10% lives in a low-density area and 11% lives in a very low-density area.

In addition, the sum score of the answers on the 7 questions from Frieling (2008) was used in the analysis. The mean total score of the self-perceived social cohesion is 22.98, with a standard deviation of 5.60. The Cronbach's Alpha of the total score of social cohesion is 0.827. The mean percentage of western ethnic minorities in the neighborhoods is 12% and also the mean percentage of non-western ethnic minorities is 12%.

Table 1: Basic sample characteristics

	Sample 2014 (%)	Sample 2008 (%)	Netherlands (%)	Eindhoven (%)
Gender				
Male	38	39	50	51
Female	62	61	50	49
Age				
Age (<40 years)	16	37	48	52
Age (40-65 years)	49	49	36	32
Age (>65 years)	35	15	16	16
Household composition				
One person household	13	11	37	39
Couple without children	45	34	29	28
Couple with children	36	47	27	21
Single parent family and other	6	7	7	12
Income				
Low income (<€3000 net/month)	54	58	71	
High income (>€3000 net/month)	41	33	29	
(Missing)	5	9		
Education				
Primary education	19	23	29	34
Secondary education	31	31	43	38
Higher education	50	46	28	27
Work				
No work	54	35	48	
Part time work	28	38	21	
Full time work	18	27	31	
Club membership				
No club membership	33	30		
One or more club memberships	67	70		

Note: N=177

4 Methods and results

A path analysis was used to analyze simultaneously the relations between the independent variables (i.e. personal and neighborhood characteristics) and the dependent variables (i.e. place attachment, social network in the neighborhood, and social satisfaction) and also the relations between the dependent variables in a single model. This analysis is an extension of the multiple regression analysis and a special case of Structural Equation Modelling (SEM). Compared to SEM, where also latent variables (i.e. unobserved variables) can be included, a path analysis only includes observed or measured variables. A path analysis can estimate simultaneously direct and indirect effects (Streiner 2005).

This section describes the results of the path analysis. The statistical software package LISREL (Jöreskog, Sörbom 2008) was used for estimating the model. The maximum likelihood method (ML) was used to estimate the models, because compared to other estimation methods, ML generally performs best (Iacobucci 2010). The maximum likelihood estimation is appropriate for small sample sizes and non-normally distributed variables (Suhr 2006).

First, based on the literature review, potentially significant relations between personal (i.e. age, gender, income, education level, health, household composition, work status, transportation modes, car ownership, home ownership, type of dwelling, and length of residence) and neighborhood characteristics (i.e. density, social cohesion level, ethnicity neighborhood, income levels in the neighborhood, average value of the dwellings, and distance to facilities), place attachment, number of interactions with local residents, size

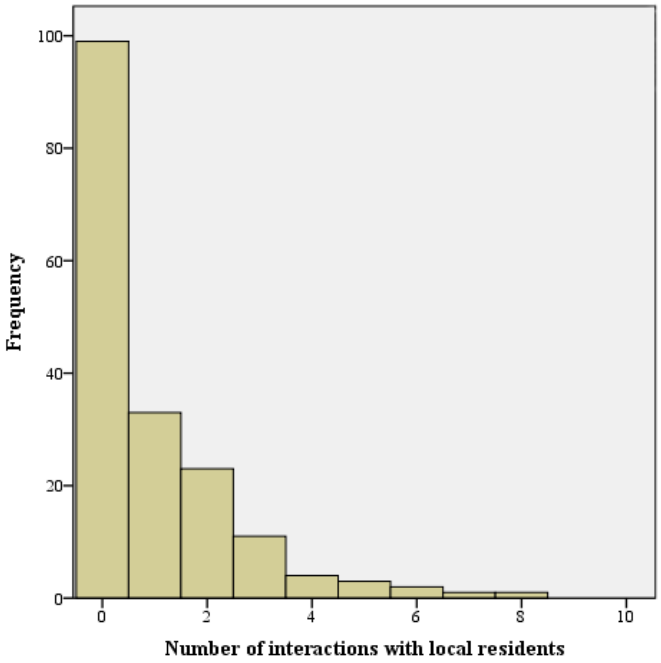


Figure 4: Distribution number of social interactions with local residents

of local social networks, and social satisfaction were identified using bivariate analyses. The relationships that were significant at the 0.10 level were then entered in the model. Second, to obtain a better model, the direct effects that were not significant at the 0.10 level in the path model were removed stepwise (e.g. education level, density levels, length of residence, home ownership, gender, or health). The variable age was also included expressed in log. However, this this did not lead to a better fitting model. Therefore, it was chosen to include age as a linear continuous variable. This resulted in a final model shown in Figure 5. Table 2 shows the mean and standard deviation of the significant variables that were included in the final model.

Table 3 shows the unstandardized coefficients and t-statistics of direct and (significant and non-significant) total effects of the final model and Table 4 shows the goodness-of-fit of the model. The model provides an adequate fit of the data, as the value of Chi Square divided by the degrees of freedom and the Normed Fit Index is close to 1 and the value of RMSEA is (close to) 0 (Golob 2003). In addition, the value of the model's Akaike information criterion (AIC) is close to the value of saturated AIC (Golob 2003). Overall, the results show a good fit of the model to the data.

4.1 The effects between endogenous variables

Place attachment was found to have a positive effect on the number of social interactions with local residents. This result suggests that people who are more attached to a neighborhood have more social interactions with local residents. This finding is in line with the literature, which suggests that place attachment affects the quality of social relationships with local residents (Dallago et al. 2009). However, no significant direct or indirect effect was found of place attachment on the number of local residents in the social network.

As can been seen in Table 3, place attachment is positively affected by the number of local residents in the social network. This result is in line with Lewicka (2010), who argued that neighborhood ties and sense of security are the social predictors of place attachment. van den Berg et al. (2014) found that place attachment is affected by the satisfaction with social contacts in the neighborhood. People are probably more satisfied with their social contacts in the neighborhood if they have more local residents in their social network. Therefore, they will probably feel more connected to their living environment.

Table 2: Variables considered in the analysis

	Mean	St. Dev.
Endogenous variables		
Social satisfaction	34.63	3.93
Place attachment	38.56	7.37
Number of local residents	3.56	4.40
Number of social interactions with local residents	0.96	1.48
Exogenous variables		
Personal characteristics		
Age	56.42	15.25
Low income (<€3000 net/month) (dummy)	0.23	0.42
Club membership (dummy)	0.67	0.47
Working (dummy)	0.46	0.50
Home ownership (dummy)	0.81	0.39
Length of residence	21.13	13.48
Car ownership (dummy)	0.90	0.30
Frequency of walking for transportation in the neighborhood	2.98	1.26
Frequency of cycling	5.63	1.88
Neighborhood characteristics		
Very high density (>2500 addresses per km^2) (dummy)	0.23	0.42
Social cohesion	22.98	5.60
% Non-western ethnic minorities	11.80	7.12

Notes: N=177

Regarding the effects of place attachment on social satisfaction, place attachment was found to have a positive effect on social satisfaction. This result implies that people who are more attached to their living environment are more satisfied with their social life. This finding is in line with results of an earlier study that showed an indirect effect of place attachment on social satisfaction, mediated by self-perceived loneliness (Weijs-Perrée et al. 2015). Overall, the results show that the attachment to the living environment plays an important role in explaining social satisfaction, and therefore the well-being of individuals.

As can been seen, no significant direct effect was found of the number of local social interactions on social satisfaction. On the other hand, an indirect positive effect was found of the number of local residents in the social network on social satisfaction, mediated by place attachment. People who have more local ties, with whom they feel somewhat or very close with, feel more attached to their neighborhood than people with less local ties. In addition, no significant direct or indirect effect was found of the number of local social interactions on social satisfaction. Weijs-Perrée et al. (2015) found a direct effect of the number of social interactions on social satisfaction. However, they focused on the total number of social interactions (e.g. friends, family, neighbors, club members, colleagues) and did not focus on specifically social interactions with local residents.

4.2 The effects of exogenous variables

Regarding the effects of personal and neighborhood characteristics on the number of local residents in the social network, the results show that frequency of walking for transportation in the neighborhood has a positive effect. This finding is in line with existing literature, which shows a relation between social interactions and frequency of walking. For example, findings of Glanz (2011) and van Cauwenberg et al. (2014) suggest that walking leads to more spontaneous social interactions. Also, frequency of cycling was found to positively affect the number of local residents in the social network. This is probably due to the fact that during cycling spontaneous social interactions with local residents are more likely to occur than when using a car as transport mode. Weijs-Perrée et al. (2015) found a similar relation between the number of social interactions and the

frequency of cycling. Very high density was found to have a negative direct effect on the number of local residents. This finding suggests that people living in a neighborhood with a very high density level have 3.11 fewer local residents in their social network than people living in a neighborhood with a high, moderate, low, or very low density. This result is in line with findings of Thomése, van Tilburg (2000) and van den Berg et al. (2011). The length of residence also showed a significant direct effect on the number of local residents in the social network. People who live for a longer time in the neighborhood probably had more time to bond with other local residents.

Regarding the effects of personal and neighborhood characteristics on the number of social interactions in the neighborhood, age was found to have a small negative effect. The estimated value of the path coefficient suggests that every year increase in age results in 0.02 social interactions with local residents per two days fewer. The average total number of social interactions in the neighborhood per two days is 1.01.This means that for every year increase in age, the number of social interactions with local residents decreases by 1.98% (3.65 social interactions). Although this is a small decrease, it suggests that elderly have fewer social interactions with local residents than younger people. In addition, home ownership was found to have significant effect on the number of social interactions in the neighborhood. This is in line with findings of previous research (e.g. Guest et al. 2006). Social cohesion, very high density, low income, car ownership, length of residence, frequency of walking, and frequency of cycling were found to have an indirect effect on the number of social interactions with local residents.

With regard to place attachment, the results show that a low income negatively affects place attachment. An explanation could be that people with low incomes often have less choice where to live and are therefore probably less attached to their living environment. In addition, people with low incomes often live in (social) rental dwellings and the attachment to the living environment of tenants is probably weaker than the attachment to the living environment of home owners, as was found by Brown et al. (2003) and van der Houwen, Kloosterman (2011). Second, social cohesion was found to have a positive direct effect on place attachment. It implies that people who live in a neighborhood with a high social cohesion level feel more attached to their neighborhood. This finding is in line with findings from other studies (e.g. Livingston et al. 2008, van den Berg et al. 2014). Finally, car ownership was found to have a significant negative effect on place attachment. People who own a car, probably are more able to maintain social contacts and to use facilities (e.g. shopping or sport facilities) at a further distance. Therefore, they probably feel less attached to their neighborhood.

Some of the personal and neighborhood characteristics were not found to have a direct effect on social satisfaction, but an indirect effect mediated by the other dependent variables. Social cohesion was found to have a significant indirect effect on social satisfaction, mediated by place attachment. Club membership is found to have a positive direct effect on social satisfaction. This implies that people who are a member of one or more clubs are more satisfied with their social life. This finding is in line with Bonsang, van Soest (2012). They concluded that non-professional activities affect the satisfaction with social contacts. Next, working is also found to have a positive direct effect on social satisfaction. This finding suggests that people who are employed are more satisfied with their social life than people who are not employed. Having a job provides social interactions with colleagues and could increase the feeling of being socially integrated. This probably explains the fact that the social satisfaction of people who have a job is higher than of people who are unemployed. These findings imply the importance of participating in social activities and being socially included for the quality of the social network. Moreover, the percentage of non-western ethnic minorities in the neighborhood is found to have a negative effect on social satisfaction. This implies that people living in neighborhood with a higher percentage of non-western ethnic minorities are less satisfied with their social life. This is in line with findings of van der Houwen, Kloosterman (2011) who suggest that a larger proportion of non-western ethnic minorities negatively affect the number of social interactions between local residents. In addition, indirect effects on social satisfaction were found of low income, frequency of cycling, frequency of walking for transportation, very high density, and the social cohesion level in the neighborhood.

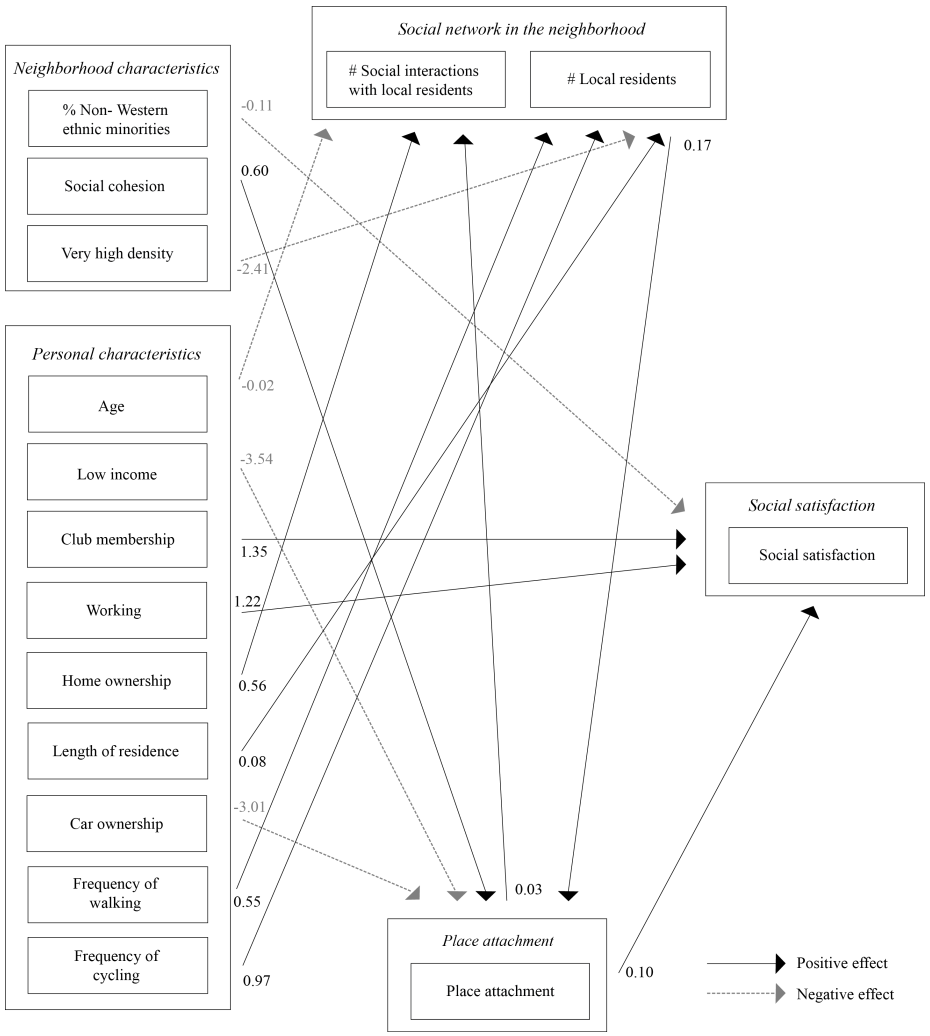


Figure 5: Estimated path model

5 Conclusion and discussion

The aim of this study is to analyze the (direct and indirect) effects of personal and neighborhood characteristics, place attachment, and social contacts in the neighborhood on social satisfaction, simultaneously in a single model. In addition, to understand the importance of the neighborhood for explaining social satisfaction, local social network characteristics and place attachment were included. Based on data collected using a social interactions diary and survey among 177 respondents in the Eindhoven region, a path model was estimated.

Social satisfaction can be considered as the quality of social life, which is therefore an indicator for an important aspect of the well-being of an individual. The estimated path model shows a good fit to the data. The model structure showed that several neighborhood and personal characteristics directly or indirectly (mediated by place attachment) influence social satisfaction. Particularly, the social dimension of the neighborhood plays a relevant role in explaining the total social satisfaction of individuals. This study provides more insight in the relations between social satisfaction, neighborhood networks, and place attachment and contributes to the knowledge gap on social satisfaction, which received still little attention in existing studies.

A previous study showed that the number of social interactions is positively related to social satisfaction (Weijs-Perrée et al. 2015). However, findings in this study show no direct relation between the number of social interactions with neighbors and social

Table 3: Path analysis model estimates

From	To							
	Social satisfaction		Place attachment		# Local residents		# Social interactions local residents	
	Direct	Total	Direct	Total	Direct	Total	Direct	Total
Effects between the endogenous variables								
Social satisfaction t statistic								
Place attachment t statistic	0.10** 2.78	0.10 2.78					0.03* 1.66	0.03 1.66
# Local residents t statistic		0.02 1.66	0.17** 2.08	0.17 2.08				
# Social interactions t statistic						0.60 2.05		
Effects of the exogenous variables								
Age t statistic							-0.02** -2.12	-0.02 -2.12
Low income t statistic		-0.35 -2.02	-3.55** -2.95	-3.55 -2.95				
Club membership t statistic	1.35** 2.37	1.35 2.37						
Working t statistic	1.17** 2.22	1.22 2.33						
Home ownership t statistic							0.56* 1.90	0.56 1.90
Car ownership t statistic			3.02* 1.77	3.02 1.77				
Length of residence t statistic					0.08** 2.61	0.08 2.61		
Frequency of walking t statistic				0.16 1.67	0.97** 2.84	0.97 2.84		
Frequency of cycling t statistic					0.55** 2.40	0.55 2.40		
Very high density t statistic					-2.42** -2.32	-2.42 -2.32		
Social cohesion t statistic		0.06 2.57	0.60** 6.73	0.60 6.73				
% Non-western ethnic minorities t statistic	-0.11** -2.93	-0.11 -2.93						
R^2	0.17		0.31		0.17		0.07	
R^2reduced	0.14		0.29		0.14		0.05	

Notes: unstandardized effects, *Significant at 0.1 level, **Significant at 0.05 level

satisfaction. This suggests that interactions with neighbors are less important than interactions with other people (e.g. family, friends, club members, colleagues) for the satisfaction of people's social life. However, it is recognized that these weak ties (i.e. neighbors) are important to discuss important matters with and to feel supported by (Small 2013, Cramm et al. 2012). Further research is needed to analyze how interactions with different weak and stronger ties could influence the total social satisfaction.

Results of this study also show that the frequencies of walking and cycling in the neighborhood have a positive effect on the share of local residents in the social network. Therefore, it is important that neighborhoods provide safe and attractive walking and cycling routes, and an attractive level of facilities that encourage cycling and walking in the neighborhood to create a healthy social environment.

Although this research showed that characteristics of the neighborhood, local ties, and place attachment are important for explaining social satisfaction, several questions remain for future research. In this study, data was collected on personal social interactions that are more than just a greeting. However, greetings are typically neighborhood social interactions; including these typical neighborhood social interactions could give more insight in the importance of local social contacts. In addition, people sometimes interact

Table 4: Goodness-of-fit statistics of the model

Statistic	Value
Degrees of Freedom	38
Minimum Fit Function Chi Square	60.35
Chi Square / Degrees of Freedom	1.59
RMSEA	0.054
Model AIC	252.28
Saturated AIC	272.00
Normed Fit Index	0.93

with neighbors only once a month or less (e.g. event in the neighborhood). Using a larger time-window of the interaction diary (i.e. more than two days), more typical interactions with neighbors could be captured. In addition, the data collection method and collecting a second wave of data among the same respondents from a study in 2008 led to a somewhat biased sample (i.e. more women, elderly, and people with a higher education level). A more representative sample of the area could increase the interpretation and generalizability of the results.

For further research, it would furthermore be interesting to analyze more detailed information on characteristics of social interactions between local residents (e.g. frequency, type, purpose, the importance of these social interactions, and location). Moreover, it would be interesting to analyze the negative effects of neighborhood social interactions on the number of social interactions outside of the neighborhood, because these interactions could limit social opportunities to outside of the neighborhood (e.g. Pinkster 2014). In addition, using a larger dataset and comparing the results of this study to other cities in the Netherlands or other countries, could give a better impression of patterns and routines of people's activities in neighborhoods. Another limitation of this research is the lack of information on residential selection which is dependent on residential preferences and restrictions (e.g. price of a dwelling). For example, people with lower incomes have more restrictions in their choice of where they want to live. This residential selection could also influence their social life. It remains, thus, a challenge for future research to avoid selection bias in research on neighborhood effects on people's social life. Furthermore, this study did not include information of the labor market (e.g. percentage of employed people in the neighborhood) or work-related characteristics (e.g. job quality). People who are more satisfied with their work life are probably also more satisfied with their social life. Including this information could give more insight into the effects of neighborhood contacts on social satisfaction, mediated by the effect of characteristics of the job status at an individual level and the labor market at the neighborhood level.

Nevertheless, findings of this study are relevant for urban planners and policymakers, who increasingly focus on improving the livability, social cohesion, and the social status of residents in urban areas. We argue that the structure of the living environment generally affects the quality of the social life of individuals. Our findings suggest that the self-perceived social cohesion level and place attachment could be increased by, for example, creating meeting spaces that encourage social interactions (e.g. green spaces), investing in safety, organizing social activities, or involving local residents in new policy plans for the neighborhood. Local residents should also be stimulated to socially mix with people from different backgrounds to increase trust levels and strengthen social cohesion in local communities. This reduces problems such as anti-social behavior, crime, and lack of trust (Randall 2012). Local policies should stimulate people to participate in social and community activities, by for example, establishing and supporting local organizations and local activities/events (Forrest, Kearns 2001).

References

Bonsang E, van Soest A (2012) Satisfaction with social contacts of older Europeans. *Social Indicators Research* 105: 273–292. CrossRef.

Brown BB, Perkins DD, Brown G (2003) Place attachment in a revitalizing neighborhood: Individual and block levels of analysis. *Journal of Environmental Psychology* 23[3]: 259–271. CrossRef.

Brueckner JK, Largey AG (2008) Social interaction and urban sprawl. *Journal of Urban Economics* 64: 18–34. CrossRef.

Buffel T, Demeere S, De Donder L, Verté D (2011) Fysieke, sociale en psychologische dimensies van de woonomgeving: Ouderen aan het woord over hun verbondenheid met de buurt. *Tijdschrift voor Sociologie* 32[1]: 59–87

Carrasco JA, Miller EJ, Wellman B (2008) How far and with whom do people socialize? Empirical evidence about the distance between social network members. *Transportation Research Record* 2076: 114–122. CrossRef.

Cattel VR (2001) Poor people, poor places, and poor health: The mediating role of social networks and social capital. *Social Science & Medicine* 52[10]: 1501–1516. CrossRef.

CBS (2012) Statline. http://www.cbs.nl (retrieved May 15, 2014)

Corti L (1993) Using diaries in social research. Social research update. University of Surrey. http://sru.soc.surrey.ac.uk/SRU2.html (retrieved May 30, 2016)

Cramm JM, van Dijk HM, Nieboer AP (2012) The importance of neighborhood social cohesion and social capital for the well-being of older adults in the community. *The Gerontologist* 53[1]: 142–152. CrossRef.

Croezen S (2010) Social relationships and healthy ageing: Epidemiological evidence for the development of a local intervention programme. PhD Thesis, Wageningen University, the Netherlands

Dallago L, Perkins DD, Santinello M, Boyce W, Molcho M, Morgan A (2009) Adolescent place attachment, social capital, and perceived safety: A comparison of 13 countries. *The American Journal of Community Psychology* 44: 148–160. CrossRef.

De Jesus M, Puleo E, Shelton RC, M EK (2010) Associations between perceived social environment and neighborhood safety: health implications. *Health & Place* 16[5]: 1007–1013. CrossRef.

Delmelle EC, Haslauer E, Prinz T (2013) Social satisfaction, commuting and neighborhoods. *The Journal of Transport Geography* 30: 110–116. CrossRef.

Farrell SJ, Aubry T, Coulombe D (2004) Neigborhoods and neighbors: Do they contribute to personal well-being. *Journal of Community Psychology* 32: 9–25. CrossRef.

Forrest R, Kearns A (2001) Social cohesion, social capital and the neighbourhood. *Urban studies* 38[12]: 2125–2143. CrossRef.

Foster KA, Pitner R, Freedman DA, Bell BA, Shaw TC (2015) Spatial dimensions of social capital. *City & Community* 14[4]: 392–409. CrossRef.

Frieling MA (2008) Een goede buur: 'joint production' als motor voor actief burgerschap in de buurt. Rijksuniversiteit Groningen, Groningen http://www.rug.nl/research/-portal/files/2727687/03_h3.pdf (retrieved January 10, 2014)

Gemeente Eindhoven (2017) Buurtmonitor Eindhoven. https://eindhoven.buurtmonitor.-nl/ (retrieved August 7, 2017)

Glanz TA (2011) Walkability, social interaction and neighborhood design. Regional planning program: Student projects and theses, paper 6. http://digitalcommons.unl.edu/arch_crp_theses/6 (retrieved February 2, 2014)

Golob T (2003) Structural equation modeling for travel behavior research. *Transportation Research Part B: Methodological* 37: 1–25. CrossRef.

Greif MJ (2009) Neighborhood attachment in the multiethnic metropolis. *City & Community* 8[1]: 27–45. CrossRef.

Guest AM, Cover JK, Matsueda RL, Kubrin CE (2006) Neighborhood context and neighboring ties. *City & Community* 5[4]: 363–385. CrossRef.

Hanibuchi T, Kondo K, Nakavaya T, Shirai K, Hirai H, Kawachi I (2012) Does walkable mean sociable? Neighbourhood determinants of social capital among older adults in Japan. *Health & Place* 18: 229–239. CrossRef.

Helliwell JF, Putnam RD (2004) The social context of well-being. *Philosophical Transactions of the Royal Society B: Biological Sciences* 359: 1435–1446. CrossRef.

Iacobucci D (2010) Structural equations modelling: Fit indices, sample size, and advanced topics. *Journal of Consumer Psychology* 20: 90–98. CrossRef.

Jöreskog KG, Sörbom D (2008) *LISREL 8 User's reference guide*. SSII, Chicago

Kawachi I, Berkman LF (2001) Social ties and mental health. *The Journal of Urban Health* 78[3]: 458–467. CrossRef.

Knipscheer CPM, De Jong-Gierveld J, Van Tilburg TG, Dykstra PA (1995) *Living arrangements and social networks of older adults*. VU University Press, Amsterdam

Kolodinsky JM, DeSisto TP, Propena D, Putnama ME, Roche E, William R (2013) It is not how far you go; it is whether you can get there: Modeling the effects of mobility on quality of life in rural New England. *Journal of Transport Geography* 31: 113–122. CrossRef.

Lansford JE, Sherman AM, Antonucci TC (1998) Satisfaction with social networks: An examination of socio-emotional selectivity theory of cohorts. *Psychology and Aging* 13: 544–552. CrossRef.

Lewicka M (2010) What makes neighborhood different from home and city? Effects of place scale on place attachment. *Journal of Environmental Psychology* 30: 35–51. CrossRef.

Livingston M, Bailey N, Kearns A (2008) People's attachment to place: The influence of neighbourhood deprivation. Project report. Charterd Institute of Housing/Joseph Rowntree Foundation, Coventry

Maas J, van Dillen SME, Verheij RA, Groenewegen PP (2009) Social contacts as a possible mechanism behind the relation between green space and health. *Health & Place* 15: 586–595. CrossRef.

Moore S, Bockenholt U, Daniel M, Frohlich K, Kestens Y, Richard L (2011) Social capital and core network ties: A validation study of individual-level social capital measures and their association with extra- and intra-neighborhood ties, and self-rated health. *Health & Place* 17[2]: 536–544. CrossRef.

Pinkster FM (2014) Neighbourhood effects as indirect effects: Evidence from a Dutch case study on the significance of neighbourhood for employment trajectories. *International Journal of Urban and Regional Research* 38[6]: 2042–2059. CrossRef.

Pinquart M, Sörensen S (2000) Influences of socioeconomic status, social network, and competence on subjective well-being in later life: A meta-analysis. *Psychology and Aging* 15[2]: 187–224. CrossRef.

Randall C (2012) Measuring national well-being – our relationships, 2012. http://-webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/-dcp171766_257045.pdf (retrieved May 30, 2016)

Rubinstein RI, Parmelee PA (1992) Attachment to place and the representation of the life course by the elderly. In: Altman I, Low SM (eds), *Human Behavior and Environment.* Plenum Press, New York, 139–163. CrossRef.

Scheffert DR, Horntvedt J, Chazdon S (2008) Social capital and our community. http://www.extension.umn.edu/community/civic-engagement/docs/social-capital-community.pdf (retrieved May 30, 2016)

Siegler V (2014) Measuring social capital. http://webarchive.nationalarchives.gov.uk/-20160105160709/http://www.ons.gov.uk/ons/dcp171766˙371693.pdf (retrieved May 30, 2016)

Sirgy MJ, Cornwell T (2002) How neighborhood features affect quality of life. *Social Indicators Research* 59[1]: 79–114. CrossRef.

Small ML (2013) Weak ties and the core discussion network: Why people regularly discuss important matters with unimportant alters. *Social networks* 35: 470–483. CrossRef.

Streiner DL (2005) Finding our way: An introduction to path analysis. *The Canadian Journal of Psychiatry* 50[2]: 115–122. CrossRef.

Suhr D (2006) The basics of structural equation modeling. Presented: Irvine, CA, SAS User Group of the Western Region of the United States (WUSS). http://www.lexjansen.-com/wuss/2006/tutorials/tut-suhr.pdf (retrieved August 8, 2017)

Taylor J, Chatters M, Hardison B, Riley A (2001) Informal social support networks and subjective well-being among African Americans. *Journal of Black Psychology* 27: 439–463. CrossRef.

Theodori GL (2001) Examining the effects of community satisfaction and attachment on individual wellbeing. *Rural Sociology* 66[4]: 618–628. CrossRef.

Thomése GCF, van Tilburg TG (2000) Neighbouring networks and environmental dependency: Differential effects of neighbourhood characteristics on the relative size and composition of neighbouring networks of older adults in the Netherlands. *Ageing & Society* 20[1]: 55–78. CrossRef.

Umberson D, Montez JK (2010) Social relationships and health: A flashpoint for health policy. *Journal of Health and Social Behavior* 51: 554–566. CrossRef.

van Cauwenberg J, de Donder L, Clarys P, de Bourdeaudhuij I, Buffel T, De Witte N, Dury S, Verté D, Deforche B (2014) Relationships between the perceived neighborhood social environment and walking for transportation among older adults. *Social Science & Medicine* 104: 23–30. CrossRef.

van den Berg PEW, Arentze TA, Timmermans HJP (2011) A path analysis of social networks, ICT use and social activity-travel patterns. Proceedings of the International Workshop Frontiers in Transportation: Social Interactions. Niagara-on-the-Lake, Canada

van den Berg PEW, Heijs WJM, Willemsen SWQ, Havermans DWQ (2014) Factors influencing place attachment and plans to move. A case study in Terneuzen, the Netherlands. Proceedings of the European Network for Housing Research (ENHR) 2014 conference 'Beyond Globalisation Remaking Housing Policy in a Complex World'. Edinburgh

van der Houwen K, Kloosterman R (2011) Vertrouwen in en contacten met buurtgenoten. https://www.cbs.nl/nl-nl/achtergrond/2011/13/vertrouwen-in-en-contacten-met-buurtgenoten, (retrieved January 10, 2014)

Vermeij L (2008) Goede schuttingen maken goede buren. In: Schnabel P, Bijl R, de Hart J (eds), *Betrekkelijke betrokkenheid. Studies in sociale cohesie Sociaal en Cultureel rapport 2008*. Sociaal en Cultureel Planbureau, Den Haag, 113–131

von Hippel W, Henry JD, Matovic D (2008) Aging and social satisfaction: Offsetting positive and negative effects. *Psychology and Aging* 23: 435–439. CrossRef.

Weijs-Perrée M, van den Berg PEW, Arentze T, Kemperman A (2015) Factors influencing social satisfaction and loneliness: A path analysis. *The Journal of Transport Geography* 45: 24–31. CrossRef.

Williams DR, Roggenbuck JW (1989) Measuring place attachment: Some preliminary results. Paper presented at the session on outdoor planning and management, NRPA Symposium on Leisure Research, San Antonio, TX

Windsor TD, Pearson EL, Crisp DA, Butterworth P, Anstey KJ (2012) *Neighbourhood characteristics: Shaping the wellbeing of older Australians*. NSPAC Research Monograph 2, National Seniors Productive Ageing Centre, Canberra

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 3, 2017, 153–173
DOI: 10.18335/region.v4i3.145

journal homepage: region.ersa.org

Regional Spanish Tourism Competitiveness

Juan Carlos Martín[1], Cira Mendoza[1], Concepción Román[1]

[1] University of Las Palmas de Gran Canaria, Las Palmas, Spain

Received: 21 July 2016/Accepted: 18 May 2017

Abstract. The aim of this paper is to analyse the regional tourist competitiveness performance in Spain. We use the seven pillars of tourism from a very detailed and complete database carried out by the Spanish Government – MoniTUR 2010 as primary data. Thus, we calculate several regional tourist competitiveness indices using data envelopment analysis (DEA) to analyse the robustness of the results obtained in the ranking of the tourist competitiveness for the 17 Spanish Autonomous Communities. Our results are robust to the use of two different modelling strategies: (1) input and output variables selection; and (2) virtual and super efficiency DEA models. Madrid and La Rioja are found to be the most competitive regions; meanwhile other inland regions of Spain like Extremadura and Aragón are the least competitive. The position of each of the laggard Autonomous Communities should be analysed by their respective destination management organizations (DMOs) in order to envisage adequate corrective measures.

Key words: DEA; Regional tourist competitiveness; Destination management organizations; MoniTUR; Virtual Efficiency

1 Introduction

Tourism has become today one of the largest and fastest-growing economic sectors worldwide. It is an important driver of socio-economic progress as it stimulates economies and leads to the creation of jobs, incomes, investments and exports. Despite the obstacles faced during the impact of the unprecedented financial and economic crisis that hit the world in 2008, the sector has proven to be a consolidated industry which still maintains high levels of activity and has contributed to economic recovery. In 2014, tourism was responsible for generating a significant 9% world GDP (including direct, indirect and induced effect). International tourist arrivals worldwide reached 1,138 million in 2014, 51 million more than in 2013, and the UNWTO forecasts a growth between 3% and 4% in 2015 (UNWTO 2015). Despite the fall in the tourism activity in 2008 and 2009 due to the destructive effects of the financial crisis, Spanish tourism has proven to recover successfully from this observed unprecedented drawback. Spain attracted in the year 2014 for the first time in history a peak of 65.2 million international tourists, getting positioned as the third world's top tourist destination in international tourist arrivals and the second in international tourism receipts. In addition, the country ranked first among 141 countries in the Travel & Tourism Competitiveness Index 2015 published by the World Economic Forum. In fact, tourism is one of the main economic driving forces in the country, contributing directly 10.9% GDP, generating 12% of jobs and covering 276% of trade deficit (Ministerio de Industria, Energía y Turismo 2015). The most significant

tourist segment in the country is sun and sand mass tourism. Cities and municipalities along the Mediterranean coast and the Balearics and the Canary Islands stand out as some of the world's sun and beach most favourite tourist destinations.

In order to be competitive, tourist destinations must persuade their potential tourists that they will obtain more benefit visiting their destination than any other else (Crouch 2011). Destination marketing plays a determinant role in this regard, assisting visitors with pre-visit information and after-arrival additional information; coordinating many constituent tourist sector elements; creating specific tourism planning laws; or helping to ensure the attractiveness of a set of events, programs and tourism facilities, among others. Nevertheless, when competences in tourism in a certain country are decentralized, regional governments cannot confer the same effort to tourism planning since the priorities and strategies for tourism might not be concordant. This is usually caused by the intrinsic and differential characteristics of territories. Thus, a regional analysis on tourist competitiveness within the Spanish territory is paramount in order to develop adequate destination marketing plans that enhance the image of the tourist Spanish brand.

Despite the favourable results of Spanish tourism, improvements in competitiveness from the regional perspective are needed in order to enhance the competitive position of the whole country and ensure a sustainable growth (Exceltur 2011b). Spain is divided into 17 regions at level II of the Nomenclature of Territorial Units for Statistics (referred as NUTS by the French acronym). These regions, also named Autonomous Communities, have regional governments that share governance with the Spanish central administration within their respective territories. In the field of tourism, the exclusive competence corresponds to the Autonomous Communities, so the medium and long term regional tourist competitive success depend largely on the regional tourist policies. Differences in tourist competitiveness among Spanish regions exist, mainly due to the existence of different territorial natural endowments. This dissimilarity makes the measurement and comparison of regional tourist competitiveness a complex and difficult task. Exceltur (2011b) reveals the existence of these differences on the base of climatic, scenic, cultural and sociodemographic characteristics that difficult the regional comparisons.

In this sense, the development of measurement techniques to benchmark regional tourist competitiveness becomes essential to boost the evolution of policies and private and public strategies that facilitate a differential gradual tourist repositioning. Alonso (2010) considers that the evaluation of competitiveness is a key aspect that allows destinations to facilitate, control and judge policies and strategies that quest for continuous improvements. The present research addresses this issue through the construction of a synthetic index to assess tourist competitiveness among the 17 Autonomous Communities in Spain. According to Cracolici et al. (2008), tourist areas endeavour to exploit their locational attractiveness by a smart use of input factors. Through Data Envelopment Analysis (DEA) methodology, multiple inputs and multiple outputs are combined to model multidimensional relationships among different regional competitiveness dimensions, and to compute several synthetic indices that measure the Spanish regional tourist competitiveness. Dealing with the seven tourism competitiveness pillars of MoniTUR 2010 report (Exceltur 2011b), the proposed approach uses DEA as a tool for multiple criteria decision making (MCDM). A further analysis of the robustness of the results obtained by different DEA methods is made and, furthermore, the paper will also analyse the potential differences observed among the Spanish regions highlighting the laggard regions.

The remainder of the paper is organized as follows: Section 2 offers some insights from the literature, section 3 describes the data section, section 4 details the methodology, section 5 presents and discusses the results, and finally section 6 offers some concluding remarks.

2 Literature review

The success of tourism destinations in world markets is determined by their competitiveness compared with other destination options (Dwyer et al. 2000). The concept of destination competitiveness can appear to be simple and easily understandable – it is the expression of

qualitative and quantitative superiority of a territory over actual and potential competitors. Nevertheless, the complexity of the concept becomes apparent when we attempt not only to define it, but also to measure it. Numerous researchers have studied the concept of destination competitiveness, their models and determinants (Scott, Lodge 1985, Crouch, Ritchie 1999, Newall 1992, OECD 1994, Dwyer, Kim 2003, Dwyer et al. 2004, Enright, Newton 2004).

Enright, Newton (2004, p. 778) state that "a destination is competitive if it can attract and satisfy potential tourists and this competitiveness is determined both by tourism-specific factors and by a much wider range of factors that influence the tourism service providers". Similarly, Ritchie, Crouch (2003, p. 2), detailed that "what makes a tourism destination truly competitive is its ability to increase tourism expenditure, to increasingly attract visitors, while providing them with satisfying, memorable experiences, and to do so in a profitable way, while enhancing the well-being of destination residents and preserving the natural capital of the destination for future generations". These authors have developed since 1990s an extensive framework that has served as key reference for tourism destination management (Crouch, Ritchie 1994, 1995, 1999, 2005, Ritchie, Crouch 1993, 2000a,b, 2003, Crouch 2011). Their proposed model has as principal strength their capability to integrate all the significant factors that might symbolize destination's tourism competitiveness. They determine in their study that the attributes explaining destination competitiveness can be gathered into four major groups: (1) Core Resources and Attractors (physiography, culture and history, market ties, mix of activities, special events, entertainment and superstructure); (2) Supporting Factors and Resources (infrastructure, accessibility, facilitating resources, hospitality, enterprise); (3) Destination Management (resources stewardship, marketing, finance and venture capital, organization, human resource development, information/research, quality of service, visitor management); and (4) Qualifying Determinants (location, interdependencies, safety/security, awareness/image/brand, cost/value) and Destination Policy, Planning and Development. Crouch (2011) analysed the impact and magnitude of these factors affecting destination competitiveness, remarking that the Core Resources and Attractors were the most relevant. The weights of the factors and sub factors were also estimated in Crouch's study, finding that Accessibility, Physiography and Climate, Positioning/Branding, Quality of Service/Experience and Safety and Security were the most significant sub factors.

The research interest for tourism destination competitiveness that emerged in the 1990s also raised concerns for measuring it. There have been a growing number of studies that measure and assess destination competitiveness, and several evaluation methods have been reported like composite indicators based on linear aggregation models or other MCDM methods like DEA or TOPSIS (The Technique for Order of Preference by Similarity to Ideal Solution). Destination competitiveness can be assessed by using both quantitative and qualitative data. The use of hard data, such as income from tourism tourist arrivals, makes possible the quantitative evaluation of destination competitiveness performance. As Kozak, Rimmington (1999) state, qualitative performance of a destination is also useful, as it ultimately drives quantitative performance. These authors measure tourist destination competitiveness for international tourist destinations by using both quantitative and qualitative approaches. In fact, Jick (1979) advocates that both methods are complements instead of substitutes. Similarly, Mendola, Volo (2017) analyze ten of the most relevant papers on the tourism destination competitiveness (TDC) topic, admitting that qualitative performance measurement would rightfully complement quantitative measures to provide guidance in tourism policy making. However, the authors do not provide a protocol or good guidelines about how to proceed with this interesting idea, and to our knowledge, most of the papers are quantitative in nature although sometimes the databases contain some qualitative or soft data.

Regarding the unit of analysis for the destinations, it can be said that there are indeed different geographical references analysing tourist competitiveness at both regional/sub regional and national levels. For example, Kayar, Kozak (2010) measure destination competitiveness for 28 European countries applying the Travel & Tourism Competitiveness Index 2007 (TTCI). There exists also a balanced presence in literature of works in

which Gooroochurn, Sugiyarto (2005) compute an aggregate index using confirmatory factor analysis to measure tourism competitiveness for over 200 countries. Dwyer, Kim (2003) built the Competitive Indicators of a Destination (CID) in order to determine the essential factors for the competitiveness of a destination. They determined the following elements: resource endowments (natural and cultural or heritage); resources created (tourist facilities, activities offered, etc.); support factors (general infrastructure, service quality, or accessibility of the destination); and destination management elements. The monitoring report on the relative tourist competitiveness of Spanish Autonomous Communities (MoniTUR) on its 2010 edition (Exceltur 2011b) analyzes their competitive and relative tourist position. It is composed by seven tourist competitive pillars, with 30 crucial areas of competitiveness and 79 final indicators. More detailed explanation of MoniTUR index may be found in the next section.

At regional and sub regional levels, Zhang et al. (2011) evaluate tourism destination competitiveness using TOPSIS and information entropy for sixteen cities in the Yangtze River Delta. The present study follows the theoretical and statistical approach of those of Cracolici, Nijkamp (2006), Cracolici et al. (2008) and Cracolici, Nijkamp (2009). These works employ DEA method in order to assess tourism competitiveness by analyzing the destination efficiency of Italian regions. Benito et al. (2014) also study the determinants of Spanish regions' tourism performance using DEA. At national level, Huang, Peng (2012) assess the competitiveness of tourism industries in nine Asian countries using Fuzzy Rasch model and TOPSIS. Abad, Kongmanwatana (2015) use DEA, super-efficiency DEA and the non-radial Nerlove-Luenberger super-efficiency DEA to compare the destination competitiveness ranking among 27 large and small countries in the European Union. Webster, Ivanov (2014) measure using a cross-section analysis the impact of competitiveness on tourism's contribution to economic growth for 131 countries.

3 MoniTUR and data

3.1 Monitoring report on the relative tourist competitiveness of Spanish Autonomous Communities (MoniTUR)

MoniTUR (Exceltur 2010, 2011b) is an index that measures the tourist relative competitiveness position of the 17 NUTS II (Autonomous Communities) in Spain. It is created by Exceltur in collaboration with Deloitte. Exceltur is a non-profit association formed by 25 of the most important companies throughout the Spanish tourism value chain and subsectors of central air, rail, sea and land transportation, lodging, travel agencies and tour operators, payment, car rental, tourist hospitals and bookings / GDS. Since 2002, these group leaders and heads of the Spanish tourism companies joined to promote a greater socio-economic recognition of what tourism, as a principal sector of Spanish economy, provides and represents and to contribute to achieve higher levels of competitiveness that consolidate a sustainable and rentable leadership and growth of tourism activity in Spain. According to Exceltur (2011b), tourism in Spain has still to face numerous challenges and opportunities with different accents and intensities according to the different Autonomous Communities. In this sense, MoniTUR 2010 is built as a benchmark to boost policy recommendations and private and public strategies in order to facilitate a gradual and differential tourist repositioning of global competitiveness and to rigorously identify the competitive risks and opportunities of each Spanish region.

MoniTUR 2010 (Exceltur 2011b) is the second and last published edition. It incorporates new areas, indicators and more accurate sources of information with respect to the previous MoniTUR 2009 edition. The MoniTUR 2010 report reveals the existence of four groups of Autonomous Communities that share structural similitudes. The first group is composed of the Community of Madrid, País Vasco and Catalonia, all with a high level of income and economic activity. A second group is distinguished by a higher specialization on littoral tourism, in which the Communities of Andalusia, the Canaries, the Balearics and Valencian Community, respectively, stand out. In fact, Catalonia, the Canaries, Balearics and Andalusia are, respectively, the main tourist Autonomous Communities. Together, they attracted 74.1% tourists who visited the country last year. The third group is comprised by mostly inland communities: Navarra, La Rioja, Castile-La Mancha

Table 1: MoniTUR 2010 global ranking

Autonomous Community	Ranking	Index. Average=100
Community of Madrid	1	114.2
Basque country	2	113.3
Catalonia	3	112.7
Andalusia	4	106.8
The Canaries	5	103.3
The Balearics	6	101.8
Valencian Community	7	101.4
Galicia	8	98.9
Navarre	9	98.8
La Rioja	10	98.7
Castile-La Mancha	11	97.8
Castile and Leon	12	97.5
Asturias	13	93.1
Murcia	14	93.0
Cantabria	15	92.1
Aragon	16	91.1
Extremadura	17	87.5

Source: Own elaboration

and Castile-and-León, together with Galicia which has coastline. Finally, a fourth group is comprised by three single-province Autonomous Communities: Murcia, Asturias and Cantabria, as well as Aragón and Extremadura which have more than one province. All of them are conditioned by a lagging tourist starting position, either by minor enhancement of their range of resources and attractions, their accessibility to markets, smaller business sector, or for their still incipient tourist planning and investments.

MoniTUR 2010 analyses 7 pillars of tourist competitiveness. The seven pillars analyses 30 determinant areas of competitiveness, which in turn are divided into 79 indicators. Table A.1 (in the appendix) shows the pillars, crucial areas for competitiveness, indicators and sub indicators that form the global index. Pillars and determinant areas are equally weighted in the index. The report identifies the competitive advantages and disadvantages for each region aggregating indicators and pillars using averages that obtain the global MoniTUR ranking (Table 1). It can be seen that the communities of Madrid, Basque Country, Catalonia, Andalusia and the Canaries were positioned as the top Autonomous Communities, respectively. The communities of Extremadura, Aragon, Cantabria, Murcia and Asturias were at the bottom of the ranking.

3.2 Data

The present study aims to analyse the tourism competitive position of Spanish NUTS II regions in Spain by composing a synthetic composite index. Data have been obtained from the MoniTUR report on its 2010 edition. MoniTUR 2010 obtained quantitative data from databases of both, public institutions, such as The National Institute of Statistics (INE), Turespaña or the Instititute of Tourism Studies (IET), and from private associations, institutes and institutions, such as Google, La Caixa, Michelin guide or the Institute of Tourism Quality of Spain. In-depth information about data resources can be found at MoniTUR 2010 annexes (Exceltur 2011a).

In the present study, the 7 pillars were chosen to be the variables used to analyse regional Spanish competitiveness. A more disaggregated analysis was discarded as the number of units is too small. The variables correspond to the average aggregation of indicators characterising each pillar. Different methods based on the selection of input and output variables as well as DEA methods will be calculated in order to evaluate the robustness of the obtained synthetic competitiveness indicators. We will compare the results of two approaches according to the consideration of variables as follows:

Table 2: Descriptive Statistics

Variables/pillars	Mean	SD	Min	Region	Max	Region
Positive factors						
Structuring and diversification of tourist products	99.83	14.12	83.38	Cantabria	141.74	Catalonia
Talent attraction, training and efficiency of human resources	100.40	16.07	75.99	the Canaries	143.08	Basque Country
Political priority and tourism governance	99.86	8.91	87.17	Cantabria	116.46	La Rioja
Performance: social and economic outcomes	99.92	15.41	81.87	Cantabria	127.33	the Canaries
Unfavourable factors						
Vision of strategic marketing and commercial support	100.94	9.17	83.50	Andalusia	115.60	Extremadura
Transport accessibility and connectivity	99.23	17.12	52.15	Madrid	120.28	Extremadura
Tourist competitive regulation and other conditions	101.29	8.93	84.27	La Rioja	114.33	Murcia

Source: Own elaboration

(1) three input variables and four output variables; and (2) one fixed input variable and the seven pillars used as output variables[1]. For the first scenario, the first three pillars were linearly transformed to become inputs. In this way, unfavourable factors or inputs are formed by 'vision of strategic marketing and commercial support', 'transport accessibility and connectivity' and 'tourist competitive regulation and other conditions'. Conversely, the variables 'structuring and diversification of tourist products', 'talent attraction, training and efficiency of human resources', 'political priority and tourism governance' and 'performance: social and economic outcomes' constitute our set of positive factors.

Table 2 details the descriptive statistics of all the pillars included in the different models according to the original values that can be regarded as outputs (second scenario). It should be noted that with the transformation into inputs, the variables 'vision of strategic marketing and commercial support', 'transport accessibility and connectivity' and 'tourist competitive regulation and other conditions' will perform better for those regions with lower values. As it can be observed, severe differences are present in the minimum and maximum values of all the pillars. Cantabria has the worst aggregated MoniTUR index in the pillars 'structuring and diversification of tourist products', 'political priority and tourism governance' and 'performance: social and economic outcomes'. However, Catalonia is the region with the highest index value of 'structuring and diversification of tourism products'. The Basque Country seems to be the region succeeding at major talent attraction, training and efficiency of human resources. La Rioja boasts the maximum index values for political priority and tourism governance and 'tourist competitive regulation and other conditions'. In addition, the Canaries are noted for its performance of social and economic outcomes, despite that it is the region with the minimum value on 'talent attraction, training and efficiency of human resources'. This is an important drawback that should be corrected in the near future if the region wants to remain competitive because, as Assaf, Josiassen (2011) contend, 'it became clear during our interviews that the service level that tourists perceive is an important determinant. Some tourists simply would not even consider visiting a country that they perceive as having a poor level of

[1]This selection is based on the rule of thumb proposed by Cooper et al. (2000). The seven used variables could be considered outputs, as they contribute positively to the measure of competitiveness. But in order to have a more discrimination power and a less restrictive constraint regarding the size of the database, some of the variables were converted to inputs with a simple normalization process.

service mindedness. The tourism industry is a perfect context for the consideration of skills in services (p. 391).' The highlights of Madrid and Andalusia are 'transport and accessibility and connectivity' and 'vision of strategic marketing and commercial support', respectively. Contrarily, Extremadura exhibits, with great differences, the minimum values on these pillars. Finally, Murcia is the region with the worst performance on 'tourist competitive regulation and other conditions'.

4 Methodology

As said, our method is based on Data Envelopment Analysis (DEA). DEA was originally designed to measure the efficiency of a firm in a context of production economics. A firm is considered efficient if it produces the maximum output for a given level of input. Charnes et al. (1978) write the seminal paper that evaluates the performance of different DMUs – decision-making units. The DMUs are characterized by being empowered entities that have the capacity of decision about how to transform inputs into outputs. This was the origin of a discipline that measures the relative efficiency of each DMU when researchers have a sample of peer observations regarding the input and output quantities (Charnes, Cooper 1985). DEA is one of the most popular non-parametric methods to assess economic efficiency whose main advantage resides in that it is possible to use multiple inputs and outputs with no prior assumption about the subjacent technology between inputs and outputs. Moreover, the method can also be applied when the sample is small (Perrigot et al. 2009).

In conventional DEA, each data component is usually classified as either an input or resource or as an output or product depending on the nature of the firm under analysis. However, Ali, Seiford (1993) argue that DEA is also an interesting method in scenarios where researchers are interested in ranking units where the existence of a production function between inputs and outputs is not so obvious. The authors concur to follow a general guideline in which the input or output classification is based according to the analysis of whether it is better or not to have lower (inputs) or higher (outputs) values. It was already explained that in our case study all the pillars from MoniTUR have the consideration of outputs, but in order to analyze to what extent the results do not depend on the conversion of outputs to inputs, two scenarios are proposed.

Ali, Seiford (1993), Charnes et al. (1994), Coelli et al. (1998), Cooper et al. (2011) and Zhu (2014) are good references to cover the basic aspects of DEA models, DEA notation, formulation and geometric interpretation. Conventional DEA models can be divided in three different classes: variable returns to scale (VRS), constant returns to scale (CRS) and additive models. These models separate the DMUs into two different sets: (1) the efficient units that lie on the frontier of the envelopment surface; and (2) those who are inefficient because they show some slack when their position is compared with the obtained frontier.

The selection of a suitable DEA model that ranks the regional tourist competitiveness index in Spain is constrained by the characteristics of the available data, the sample size and the intrinsic nature of the issue under analysis. In particular, as in our case, if researchers are interested in obtaining a full rank of all the regions under analysis, then conventional DEA models might not be appropriate, and then some new refinement of DEA needs to be applied.

In DEA analysis, it is generally assumed that there are n production units to be evaluated, using amounts of m different inputs to produce quantities of s different outputs. Specifically, the o^{th} production unit consumes x_{io} units of input i $(i = 1 \dots m)$ and produces y_{ro} units of output r $(r = 1 \dots s)$. The o^{th} production unit can now be described more compactly with the vector (X_o, Y_o), which denote, respectively, the vectors of input and output values for DMU_o.

Next, it is necessary to determine a potential set of possible dominant or non-dominant comparisons for each production unit considered in the analysis. DEA usually considers the dominance of all the possible linear combinations of the n DMUs, i.e. $(\sum_k \lambda_k X_k, \sum_k \lambda_k Y_k)$, with the scalar restricted to be non-negative[2]. The production unit o is dominated, in

[2]Different envelopment surfaces may be obtained considering additional constraints about the scalars.

terms of inputs, if at least one linear combination of production units shows that some input can be decreased without making the rest of inputs and outputs worse off. If at least one linear combination of production units shows that some output can be increased without negatively affecting the rest of inputs and outputs[3], it is dominated in terms of outputs.

In our case, policy makers and DMO managers can affect the regional tourist competitiveness by planning and implementing policies that improve the performance for some of the indicators and sub indicators included in the MoniTUR. For example, it would not be difficult to implement a policy that increases the tourist product diversification contemplating and including more tourist segments. In this paper, and given the nature of the issue under analysis, the robustness of the results is going to be based on different DEA methods based on: (1) virtual constant and variable returns to scale with an output orientation model; and (2) super-efficiency DEA model.

Formally, the multiplier-DEA CRS output efficiency for the unit o is calculated through the following linear programming problem:

$$\begin{aligned} & \min_{\nu,\mu} \frac{\sum_{i=1}^{m} \nu_i x_{io}}{\sum_{r=1}^{s} \mu_r y_{ro}} \\ s.t. \quad & \frac{\sum_{i=1}^{m} \nu_i x_{ij}}{\sum_{r=1}^{s} \mu_r y_{rj}} \geq 1 \ (j = 1 \ldots n), \\ \text{where} \quad & \nu_i, \mu_r \geq 0 \end{aligned} \tag{1}$$

The set of constraints requires that the same weights, when applied to all the countries, do not provide any region with efficiency lower than one. The solution to this minimization problem is not unique. Coelli (1996) shows that if there exists a solution (ν, μ) to the above problem, then there exists an infinite number of solutions because $(\phi\nu, \phi\mu)$ with $\phi \geq 0$ is also a solution to the problem. For this reason, the problem is reformulated in an equivalent linear programming program as follows:

$$\begin{aligned} & \min_{\nu,\mu} \sum_{i=1}^{m} \nu_i x_{io} \\ s.t. \quad & \sum_{i=1}^{m} \nu_i x_{ij} - \sum_{r=1}^{s} \mu_r y_{rj} \geq 0 \ (j = 1 \ldots n), \\ & \sum_{r=1}^{s} \mu_r y_{ro} = 1 \\ \text{where} \quad & \nu_i, \mu_r \geq 0 \end{aligned} \tag{2}$$

A region o is in the frontier if and only if $\sum_{i=1}^{m} \nu_i x_{io} = 1$ in optimality. The constraint $\sum_{r=1}^{s} \mu_r y_{ro} = 1$ is known as a normalization constraint, and the weighted input and output are called virtual input and virtual output, respectively (Seiford, Thrall 1990). The efficiency ratio ranges from 1 to infinity, and each region under analysis obtains their optimal multipliers minimizing the self-efficiency, given the constraints.

As explained above, this method needs a major refinement as we are interested in ranking all the regions in Spain, and this is not possible for those that lie on the frontier. The discrimination power is even more acute in our case as we do not have a large sample size. For this reason, the analysis is constrained within the pillars of the database. Our database does not even satisfy the standard rule of thumb proposed by Cooper et al.

For example, variable returns to scale models (VRS) are obtained imposing that the sum of scalars is equal to one; and non-increasing return to scale models (NIRS) are characterized by the restriction of the sum of scalars being less or equal to one.

[3] This discussion is very close to the definition of Pareto-Koopmans efficiency. The unit o is considered fully efficient if and only if the performance of other DMUs does not provide evidence that some of the inputs or outputs of the unit o could have been improved without worsening off some of its other inputs or outputs. This definition of relative performance has its origin in Farrel (1957).

(2000): $n \geq \max(m \times s, 3(m+s))$, where n is the number of DMUs, m is the number of inputs and s in the number of outputs.

Ranking DEA methods have been analyzed and proposed in previous studies. There are different techniques and proposals based on: (1) cross-efficiency DEA models (Sexton et al. 1986, Doyle, Green 1994, Adler et al. 2002); (2) super-efficiency models (Andersen, Petersen 1993, Zhu 1996, Seiford, Zhu 1998); and (3) virtual efficiency models (Bazargan, Vasigh 2003, Martín, Román 2006, 2007, Barzegarinegad et al. 2014, Martín et al. 2017).

This paper uses two different approaches based on the super-efficiency DEA models and the virtual efficiency models. Super-efficiency DEA models are based on the evaluation of a DMU when this is not included in the reference set of the envelopment models. The super-efficiency models are thus calculated using the technological frontier constructed from the rest of the DMUs included in the analysis. Super-efficiency models are also very convenient to analyze the performance of the extreme efficient units.

On the other hand, the virtual efficiency models are based on the introduction of a virtual super-efficient region in the dataset. This method discriminates all the units as there is only one efficient region (this virtual champion). The efficient frontier, based on this model, therefore consists of only this virtual super-efficient region that has been constructed ad-hoc. The ranking is justified because the same virtual unit is used for all DMUs as the reference set.

The input and output vectors for this virtual super-efficient unit are:

$$\begin{aligned} X_\nu &= \min_j\{X_j\} \\ Y_\nu &= \max_j\{Y_j\} \end{aligned}$$

where X_ν and Y_ν are the input and output vectors of the virtual super-efficient unit and X_j and Y_j are the input–output vectors of the j^{th} region. In other words, the virtual region has the lowest input vector and the highest output vector of all the regions conserved in the analysis. Thus, the DEA model expressed by (2) is run with the inclusion of this new virtual unit and the efficiency scores are used to fully rank the Spanish regional tourist competitiveness. As expected by construction, the discrimination power of the method is maximum.

5 Results

Table 3 shows the results for the 17 Autonomous Communities of Spain under the two different scenarios considered regarding the selection of input and output variables and the three selected DEA models, the virtual efficient DEA model under constant and variable returns to scale and the super efficiency model under variable returns to scale. The first column shows the results of the virtual efficiency under constant returns to scale, and it can be seen that there are two regions that can be considered equally competitive: Madrid and La Rioja. An analysis of the table reveals that, according to this methodology, the five most competitive regions are Madrid, La Rioja, the Basque country, Galicia and Andalusia. In regard to the 5 least competitive regions (Aragon, Murcia, Extremadura, Cantabria and Castile and Leon) it can be seen that, with the exception of Murcia, the regions are not located in the Mediterranean coast.

This proposal seems to penalize some of the competitive regions according to the average figures from MoniTUR. In particular, the relative positions of the Balearic archipelago and the Valencian community shifts their ranking positions very much. The shift in position of La Rioja is also particularly relevant. According to MoniTUR, La Rioja occupies the tenth position, however, from our analysis it can now be considered one of the most competitive tourist regions in Spain with Madrid.

Looking at the tail of the distribution, it can be observed that the set of the four least competitive regions is not altered within both methods: MoniTUR and VDEA-CRS rankings. But, the relative position of the regions change. If we analyse now the ranking obtained by the model under variable returns to scale, it can be concluded that the most competitive regions in Spain are again Madrid and La Rioja, but there are now two

Table 3: DEA-MoniTUR index. Global ranking

Autonomous Community	Ranking	DEA Index
Community of Madrid	1	1.000
Basque country	3	1.037
Catalonia	6	1.089
Andalusia	5	1.080
The Canaries	7	1.165
The Balearics	12	1.353
Valencian Community	10	1.326
Galicia	4	1.068
Navarre	8	1.195
La Rioja	1	1.000
Castile-La Mancha	11	1.336
Castile and Leon	13	1.461
Asturias	9	1.217
Murcia	16	1.724
Cantabria	14	1.518
Aragon	17	1.732
Extremadura	15	1.635

Source: Own elaboration

additional communities that can be considered equally competitive: Catalonia and the Basque Country. The case of Catalonia is particularly relevant as it gains four positions in the ranking. Looking at the tail of the distribution, it can be seen that the position of the five laggard regions has changed but the set of the regions remains stable: Castile-León, Cantabria, Extremadura, Murcia and Aragón. The super-efficiency model in the first scenario that uses three pillars like inputs and the rest of the pillars as outputs is not very informative as there are six regions that belong to the set of extreme efficient units for which the linear program is unfeasible. Nevertheless, regarding the five least competitive regions, it can be observed that the set is also the same.

Analysing now the results for the second scenario, it can be seen that there is not any difference between the models based on virtual efficiency independently of what type of returns to scale is used. It is not a surprise that the most competitive regions according to these models are almost the same as those that exhibit an unfeasible solution for the first scenario under the super-efficiency model. It is less informative than other models as there are five competitive regions in Spain: Madrid, the Basque Country, Catalonia, Andalusia and La Rioja. Castile-La Mancha, Castile-León, Murcia, Aragón and Extremadura are the laggard regions in Spain with respect to tourist competitiveness. Finally, analysing the results obtained for the second scenario and the super-efficiency model, it can be seen that the model ranks all the Spanish regions, being the five most competitive: (1) Madrid; (2) Catalonia; (3) the Basque Country; (4) La Rioja; and (5) Andalusia. On the opposite side, the five least competitive regions are: (13) Castile-León; (14) Cantabria; (15) Murcia; (16) Extremadura; and (17) Aragón.

The position of Madrid is not surprising as it contains the capital city of the country, and its relative ranking in the seven pillars is always well positioned. In particular, Madrid shows the best performance in the pillar of transport accessibility and connectivity. Madrid has also the most impressive art museums in Spain, being the Prado, Reina Sofia and Thissen museums. On top of being some of the most important art museums in the world, they are located in an area where all of them can be easily reached by public transport from any neighbourhood of Madrid. Madrid has also three important World Heritage sites: El Escorial, Aranjuez and Alcalá de Henares. Ortega-Martínez, Such-Devesa (2013) find that Spanish National Tourist Organization (Turespaña) and Madrid websites that promote the tourism in Madrid do not use the denomination of World Heritage in promoting and communicating potential attractions in the community of Madrid. In the specific websites of the municipalities or even the main attractions, the World Heritage distinctive plays a

more determinant role, although there are important differences between them. In 2014, the top ten visited places in Madrid were: (1) the Reina Sofia museum; (2) the Prado museum; (3) the Warner Bros Park; (4) the Amusement Park; (5) the Royal Palace; (6) the Thyssen museum; (7) the Real Madrid Santiago Bernabeu Tour; (8) the Zoo and Aquarium; (9) the lighted house; and (10) the Archaeological museum.

Another interesting result to highlight is that La Rioja shifts dramatically the position between different indices, the MoniTUR and all the indices proposed in this study. La Rioja always increases some positions from the tenth to the set of the most competitive regions in Spain. It can be seen that La Rioja shows the best performance in two out of the seven pillars: (1) the political priority and tourist governance; and (2) the tourist competitive regulation and other conditions. La Rioja is one of the inland communities in Spain that was more proactive in the past to boost tourist activities as a way to increase the value of its natural and cultural heritage (Vera, Marchena 1998, Ivars Baidal 2004). La Rioja also benefitted from the structural funds provided by the European Union, as a tool for stabilizing the rural population, that include concrete measures related to the promotion of tourism investment in rural areas through mainly agro-tourism, wine-tourism and rural tourism (Diéguez et al. 2014). In La Rioja, the importance of the protected denomination of origin (PDO) of "rioja wine" can be considered one of the most important causes that has fostered the development of wine tourism in this inland Spanish region, which has generated one of the most successful tourism products, creating important synergies with other resources such as gastronomy, monuments and cultural events (López-Guzmán, Sánchez Cañizares 2008). Cabello, Pascual Bellido (2015) contend that wine tourism has become the hallmark of the region for investments and support, both from the public and private sectors. The main Rioja wineries follow the Bordeaux wineries model, seeing in tourism a new opportunity to diversify its product, conditioning and opening not only the cellars but the vineyards to the tourist experience. The offer is enlarged with other activities such as accommodation, catering, education, training and spa, linking all of them to a broad concept of wine culture.

La Rioja competitiveness performance goes further beyond the definition provided by Ritchie, Crouch (2003, p. 2): "What makes a tourism destination truly competitive is its ability to increasingly attract visitors". It can be seen that this definition can be mainly approximated by the seventh pillar: 'the social and economic performance'. It is not strange that analysing the group of peers in which Extremadura should base the strategy in order to be more competitive, La Rioja plays a prominent role. In this regard, Extremadura can learn from the past experience of La Rioja in order to develop an adequate strategy with respect to other agricultural products that could become important icons for future agro-tourism development.

6 Conclusions

This paper presents a comparison of the MoniTUR index with six additional indices based on two different data scenarios and three methodological proposals, virtual efficiency with constant and variable returns to scale and super-efficiency. All these indices measure the Spanish regional tourist competitiveness for the year 2010. Our analysis is based on the seven pillars that conform the MoniTUR database. As it can be read in Exceltur (2011b), MoniTUR aims to provide an objective instrument that periodically assesses and compares the main transversal tourist competitiveness pillars among the seventeen Spanish Autonomous Communities.

Our results show that the proposed indices are quite robust regarding the sets of competitive and uncompetitive regions. It also seems that the indices based on DEA are particularly affecting the relative positions among three particular regions: the Balearic Islands, the Valencian Community and La Rioja. The two first regions are well-known tourist destinations where sun-and-sea tourism is one of the main segments that has been developed and attended. The beaches and good weather is one of the leading reasons why many tourists still spend holidays at destinations. Barros et al. (2011) and Benito et al. (2014) find that coastal regions are more competitive than inland regions using two-stage DEA models to analyse the regional tourist competitiveness in Spain and France. On the

other hand, La Rioja developed a strategy to promote a greater cooperation between the private and public sector that took advantage of the wine PDO and positioned the region as the best place in Spain to enjoy an authentic experience of the culture of wine.

Our main contribution is that while MoniTUR measures tourist regional competitiveness in Spain with the average figures for the seven pillars included in the analysis, our method ranks the competitive performance using very different alternatives with pertinent robust results. It is not the purpose of this paper to elicit the best proposal but as the results are robust it can be concluded that the virtual efficiency CRS model for the first scenario and the super-efficiency model for the second scenario are two valid methods to analyse the Spanish tourist regional competitiveness. From here, this study can be used by all the main stakeholders that are part of Exceltur with a clear objective for highlighting areas for resource allocation and future investments that improve the relative regional competitiveness. Other stakeholders who can be benefitted from these results are those policy planners from the laggard regions.

On the other hand, an important venue for future research should be based on having a better understanding of the relevant attributes that determine the relative tourist regional competitiveness, with special emphasis in turning the focus of research more toward assessing the relative importance of these attributes (Crouch 2011).

Acknowledgement

This research was supported by a project "La calidad del servicio en la industria hotelera. ECO2011-23852" funded by the Ministry of Science and Innovation of the Spanish Government.

References

Abad A, Kongmanwatana P (2015) Comparison of destination competitiveness ranking in the European Union using a non-parametric approach. *Tourism Economics* 21[2]: 267–281. CrossRef.

Adler N, Friedman L, Sinuany-Stern Z (2002) Review of ranking methods in the data envelopment analysis context. *European Journal of Operational Research* 140: 249–265. CrossRef.

Ali A, Seiford LM (1993) The mathematical programming approach to efficiency analysis. In: Fried HO, Lovell CAK, Schmidt SS (eds), *The Measurement of Productive Efficiency: Techniques and Applications*. Oxford University Press, New York, 120–159. CrossRef.

Alonso V (2010) Factores críticos de éxito y evaluación de la competitividad de los destinos turístico. *Estudios y perspectivas en turismo* 19[2]: 201–220

Andersen P, Petersen NC (1993) A procedure for ranking efficient units in data envelopment analysis. *Management Science* 39: 1261–1294. CrossRef.

Assaf AG, Josiassen A (2011) Identifying and ranking the determinants of tourism performance: A global investigation. *Journal of Travel Research* 51[4]: 388–399. CrossRef.

Barros CP, Botti L, Peypoch N, Robinot E, Solonandrasana B, Assaf AG (2011) Performance of French destinations: Tourism attraction perspectives. *Tourism Management* 32: 141–146. CrossRef.

Barzegarinegad A, Jahanshahloo G, Rostamy-Malkhalifeh M (2014) A full ranking for decision making units using ideal and anti-ideal points in DEA. *The Scientific World Journal*. CrossRef.

Bazargan M, Vasigh B (2003) Size versus efficiency: A case study of US commercial airports. *Journal of Air Transport Management* 9: 187–193. CrossRef.

Benito B, Solana J, López P (2014) Determinants of spanish regions' tourism performance: A two-stage, double-bootstrap data envelopment analysis. *Tourism Economics* 20[5]: 987–1012. CrossRef.

Cabello SA, Pascual Bellido N (2015) La construcción del turismo en nuevos destinos: luces y sombras. El caso de La Rioja (España). Nóesis. *Revista de Ciencias Sociales y Humanidades* 24[47-1]: 30–48. CrossRef.

Charnes A, Cooper W, Lewin AY, Seiford LM (1994) *Data Envelopment Analysis. Theory, Methodology and Applications*. Kluwer Academic, Boston. CrossRef.

Charnes A, Cooper WW (1985) Preface to topics in data envelopment analysis. *Annals of Operations Research* 2[1]: 59–94. CrossRef.

Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research* 2[6]: 429–444. CrossRef.

Coelli T (1996) A guide to DEAP version 2.1: a data envelopment analysis (computer) program. CEPA Working Paper 96/08. Centre for Efficiency and Productivity Analysis. Armidale: University of New England

Coelli T, Rao DSP, Battese GE (1998) *An Introduction to Efficiency and Productivity Analysis*. Kluwer Academic, Boston. CrossRef.

Cooper W, Sieford L, Tone K (2000) *Data Envelopment Analysis. A Comprehensive Text with Models, Applications, Reference and DEA-Solver software*. Kluwer Academic Publishers, Norwell

Cooper WW, Seiford LM, Zhu J (2011) *Handbook on data envelopment analysis*. Springer Science & Business Media, New York. CrossRef.

Cracolici MF, Nijkamp P (2006) Competition among tourist destination: An application of data envelopment analysis to Italian provinces. In: Giaoutzi M, Nijkamp P (eds), *Tourism and Regional Development. New Pathways*. Ashgate, Aldershot, England, 133–152

Cracolici MF, Nijkamp P (2009) The attractiveness and competitiveness of tourist destinations: A study of Southern Italian regions. *Tourism Management* 30[3]: 336–344. CrossRef.

Cracolici MF, Nijkamp P, Rietveld P (2008) Assessment of tourism competitiveness by analysing destination efficiency. *Tourism Economics* 14[2]: 325–342. CrossRef.

Crouch GI (2011) Destination competitiveness: An analysis of determinant attributes. *Journal of Travel Research* 50[1]: 27–45. CrossRef.

Crouch GI, Ritchie JRB (1994) Destination competitiveness: Exploring foundations for a long-term research program. Proceedings of the Administrative Sciences Association of Canada 1994 annual conference, Halifax, Nova Scotia, 79-88

Crouch GI, Ritchie JRB (1995) Destination competitiveness and the role of the tourism enterprise. Proceedings of the Fourth Annual World Business Congress, Istanbul, Turkey, 43-48

Crouch GI, Ritchie JRB (1999) Tourism, competitiveness, and societal prosperity. *Journal of Business Research* 44: 137–152. CrossRef.

Crouch GI, Ritchie JRB (2005) Application of the analytic hierarchy process to tourism choice and decision making: A review and illustration applied to destination competitiveness. *Tourism Analysis* 10: 17–25. CrossRef.

Diéguez MI, Sinde AI, Gueimonde A (2014) Actividad empresarial y resultados percebidos más allá de la racionalidad económica: El caso del turismo rural. *PASOS. Revista de Turismo y Patrimonio Cultural* 12[1]: 79–93. CrossRef.

Doyle JR, Green R (1994) Efficiency and cross-efficiency in data envelopment analysis: Derivatives, meanings and uses. *Journal of the Operational Research Society* 45[5]: 567–578. CrossRef.

Dwyer L, Forsyth P, Rao P (2000) The price competitiveness of travel and tourism: A comparison of 19 destinations. *Tourism Management* 21[1]: 9–22. CrossRef.

Dwyer L, Kim C (2003) Destination competitiveness: Determinants and indicators. *Current Issues in Tourism* 6[5]: 369–414. CrossRef.

Dwyer L, Mellor R, Livaic Z, Edwards D, Kim C (2004) Attributes of destination competitiveness: A factor analysis. *Tourism Analysis* 9[1-2]: 91–10. CrossRef.

Enright MJ, Newton J (2004) Tourism destination competitiveness: A quantitative approach. *Tourism Management* 25[6]: 777–788. CrossRef.

Exceltur (2010) Monitur 2009. Monitor de competitividad turística relativa de las comunidades autónomas españolas. available at http://exceltur.org/wp-content/uploads/-2014/10/MONITUR-2009_INFORME.pdf. (accessed 21st December 2015)

Exceltur (2011a) Monitur 2010. anexos. available at http://exceltur.org/wp-content/uploads/2014.old/10/MONITUR-2010_ANEXOS.pdf. (accessed 20th December 2015)

Exceltur (2011b) Monitur 2010. Monitor de competitividad turística relativa de las comunidades autónomas españolas. available at http://exceltur.org/wp-content/uploads/2014.old/10/MONITUR-2010_INFORME.pdf. (accessed 20th December 2015)

Farrel MJ (1957) The measurement of productive efficiency. *Journal of the Royal Statistical Society* A120: 253–290. CrossRef.

Gooroochurn N, Sugiyarto G (2005) Competitiveness indicators in the travel and tourism industry. *Tourism Economics* 11[1]: 25–43. CrossRef.

Huang JH, Peng KH (2012) Fuzzy Rasch model in TOPSIS: A new approach for generating fuzzy numbers to assess the competitiveness of the tourism industries in Asian countries. *Tourism Management* 33[2]: 456–465. CrossRef.

Ivars Baidal JA (2004) Regional tourism planning in Spain: Evolution and perspectives. *Annals of Tourism Research* 31[2]: 313–333. CrossRef.

Jick TD (1979) Mixing qualitative and quantitative methods: Triangulation in action. *Administrative science quarterly* 24[4]: 602–611. CrossRef.

Kayar cH, Kozak N (2010) Measuring destination competitiveness: An application of the travel and tourism competitiveness index (2007). *Journal of Hospitality Marketing & Management* 19[3]: 203–216. CrossRef.

Kozak M, Rimmington M (1999) Measuring tourist destination competitiveness: Conceptual considerations and empirical findings. *International Journal of Hospitality Management* 18[3]: 273–283. CrossRef.

López-Guzmán TJ, Sánchez Cañizares SM (2008) La creación de productos turísticos utilizando rutas enológicas. *PASOS. Revista de Turismo y Patrimonio Cultural* 6[2]: 159–171. CrossRef.

Martín JC, Mendoza C, Román C (2017) A DEA travel-tourism competitiveness index. *Social Indicators Research* 130[3]: 937–957. CrossRef.

Martín JC, Román C (2006) A benchmarking analysis of Spanish commercial airports. A comparison between SMOP and DEA ranking methods. *Networks and Spatial Economics* 6[2]: 111–134. CrossRef.

Martín JC, Román C (2007) Political opportunists and mavericks? A typology of Spanish airports. *International Journal of Transport Economics* 34[2]: 245–269

Mendola D, Volo S (2017) Building composite indicators in tourism studies: Measurements and applications in tourism destination competitiveness. *Tourism Management* 59: 541–553. CrossRef.

Ministerio de Industria, Energía y Turismo (2015) Balance año turístico 2014. available at http://www.minetur.gob.es/es-es/gabineteprensa/notasprensa/2015/documents/-150121dossier%20balance%20tur%C3%ADstico%202014.pdf. (accessed 3rd November 2015)

Newall JE (1992) The challenge of competitiveness. *Business quarterly* 56[4]: 94–100

OECD (1994) The world competitiveness report. World Economic Forum and IMD International, Lausanne, Switzerland

Ortega-Martínez E, Such-Devesa MJ (2013) Comunicación y conocimiento del patrimonio mundial de la Comunidad de Madrid. *Cuadernos de turismo* 31: 263–288

Perrigot R, Cliquet G, Piot-Lepetit I (2009) Plural form chain and efficiency: Insights from French hotel chains and DEA methodology. *European Management Journal* 27: 268–280. CrossRef.

Ritchie JRB, Crouch GI (1993) Competitiveness in international tourism: A framework for understanding and analysis. Proceedings of the 43rd Congress of the Association Internationale d'Experts Scientifique du Tourisme, San Carlos de Bariloche, Argentina, 23-71

Ritchie JRB, Crouch GI (2000a) Are destination stars born or made: Must a competitive destination have star genes? Proceedings of the 31st Annual Travel and Tourism Research Association Conference, Burbank, CA, 306-315

Ritchie JRB, Crouch GI (2000b) The competitive destination: A sustainability perspective. *Tourism Management* 21: 1–7

Ritchie JRB, Crouch GI (2003) *The competitive destination: A sustainable tourism perspective*. CABI, Wallingford. CrossRef.

Scott BR, Lodge GC (1985) *U.S. competitiveness in the World Economy*. Harvard Business School Press, Boston

Seiford LM, Thrall RM (1990) Recent developments in data envelopment analysis: The mathematical programming approach to frontier analysis. *Journal of Econometrics* 46: 7–38. CrossRef.

Seiford LM, Zhu J (1998) Stability regions for maintaining efficiency in data envelopment analysis. *European Journal of Operational Research* 108: 127–139. CrossRef.

Sexton TR, Silkman RH, Hogan AJ (1986) Data envelopment analysis: Critique and extensions. In: Silkman RH (ed), *Measuring Efficiency: An Assessment of Data Envelopment Analysis*. Jossey-Bass, San Francisco, 73–105

UNWTO – World Tourism Organization (2015) Tourism highlights. 2015 edition. available at http://www.e-unwto.org/doi/pdf/10.18111/9789284416899. (accessed 30th October 2015)

Vera F, Marchena M (1998) Efectos del turismo en las estructuras regionales periféricas. Una aproximación analítica. *Millars* 21: 109–144

Webster C, Ivanov S (2014) Transforming competitiveness into economic benefits: Does tourism stimulate economic growth in more competitive destinations? *Tourism Management* 40: 137–140. CrossRef.

Zhang H, Gu CL, Gu LW, Zhang Y (2011) The evaluation of tourism destination competitiveness by TOPSIS & information entropy – A case in the Yangtze river delta of China. *Tourism Management* 32[2]: 443–451. CrossRef.

Zhu J (1996) Robustness of the efficient DMUs in data envelopment analysis. *European Journal of Operational Research* 90: 451–460. CrossRef.

Zhu J (2014) *Quantitative models for performance evaluation and benchmarking: Data envelopment analysis with spreadsheets*. Springer, New York

A Appendix

Table A.1: MoniTUR 2010 tourism competitive pillars, crucial areas, indicators and sub indicators

Pillars	Crucial areas for competitiveness	Indicators	Sub indicators
1. Vision of strategic marketing and commercial support	1.1 Marketing strategies	1.1.a Strategic focus and consistency of institutional Tourism Marketing Plan	
		1.1.b Perceived efficiency of marketing strategy and institutional promotion	
	1.2 Budget allocations for marketing and promotion	1.2 Budget allocations for marketing and promotion	
	1.3 Online marketing strategy	1.3.a Commercial vocation of institutional tourism website	
		1.3.b Language accessibility of tourism institutional website	
		1.3.c Marketing on search engines	1.3.c.1 Website average position on search engines
			1.3.c.2 Results of active marketing strategies on search engines
		1.3.d Marketing on social networks and related media	1.3.d.1 Social networks positioning
			1.3.d.2 Proactive capacity on multimedia tools
	1.4 Institutional marketing management	1.4 Institutional marketing management	
2. Transport accessibility and connectivity	2.1 Air connectivity and accessibility	2.1.a Flight proficiency	
		2.1.b Air connectivity	
	2.2 Rail accessibility and connectivity	2.2.a Sufficiency of high-end trains	
		2.2.b Rail connectivity	
	2.3 Overland accessibility and	2.3 Overland connectivity	

Continued on next page

Table A.1 – continued from previous page

Pillars	Crucial areas for competitiveness	Indicators	Sub indicators
	connectivity		
3. Tourist competitive regulation and other conditions	3.1 Territory protection	3.1 Territory protection	
	3.2 Urban density on destinations	3.2.a Land occupancy on tourism destinations	
		3.2.b Congestion level of tourism destinations	
	3.3 Attractiveness of public spaces	3.3 Attractiveness of public spaces	
	3.4 Environmental Commitment	3.4.a Waste treatment	
		3.4.b Water purification	
		3.4.c Environmental manage of beaches	
		3.4.d Participation and/or membership to programs for environmental commitment	
	3.5 Other support services	3.5.a Provision of health services	3.5.a.1 Provision of health services. Doctors
			3.5.a.2 Provision of health services. Beds
		3.5.b Security levels	
4. Structuring and diversification of tourist products	4.1 Integral management of the supply by product clubs	4.1.a Clubs of products developed by Autonomous Communities	
		4.1.b Clubs of products developed together with other entities	
	4.2 Diversification of products supply	4.2.a Beach tourism	4.2.a.1 Enhancement of beaches with equipment
			4.2.a.2 Beaches with certified quality
		4.2.b Cultural tourism	4.2.b.1 Real heritage
			4.2.b.2 Assets of cultural interest
			4.2.b.3 Number of visitors to museums
		4.2.c Meeting tourism	4.2.c.1 Capacity for meetings

Continued on next page

Table A.1 – continued from previous page

Pillars	Crucial areas for competitiveness	Indicators	Sub indicators
			4.2.c.2 Number of meeting attendees
		4.2.d Nature tourism	4.2.d.1 Enhancement of natural parks
			4.2.d.2 Development of greenways
			4.2.d.3 Endowment of approved trails
		4.2.e Golf tourism	
		4.2.f Cruise tourism	
		4.2.g Nautical tourism	
		4.2.h Snow tourism	
		4.2.i Spa tourism	
		4.2.j Language tourism	4.2.j.1 Accredited Spanish teaching centres
			4.2.j.2 Spanish students
		4.2.k Gastronomic tourism	
		4.2.l Rural tourism	4.2.l.1 Provision of accommodation beds
			4.2.l.2 Effective accommodation demand
		4.2.m Wine tourism	4.2.m.1 Provision of wineries in wine tours
			4.2.m.2 Visitors in wineries in wine tours
		4.2.n Shopping tourism	4.2.n.1 Opening hours
			4.2.n.2 Commercial provisioning
	4.3 Qualification of the lodging and restaurants supply	4.3.a Commitment for formal offer	
		4.3.b Presence of upscale hotels	
		4.3.c Quality of restoration	
5. Talent attraction,	5.1 Workers productivity	5.1 Workers productivity	

Continued on next page

Table A.1 – continued from previous page

Pillars	Crucial areas for competitiveness	Indicators	Sub indicators
training and efficiency of human resources	5.2 Human capital endowment of tourist workers 5.3 Employment stability 5.4 Quality of the training system	5.2 Human capital endowment of tourist workers 5.3 Employment stability 5.4.a Attractiveness of tourist university supply 5.4.b Quality of vocational training 5.4.c Support to lifelong learning	
6. Political priority and tourism governance	6.1 Political priority 6.2 Strategic vision 6.3 Tracking of economic performance and commitment to innovation 6.4 Institutionalization of cooperation with the private sector 6.5 Effectiveness of management of tourism competences 6.6 Inter-administrative coordination	6.1.a Position of tourism in the organization level 6.1.b Budgetary effort 6.1.c Inter-councils coordination 6.2 Strategic vision 6.3.a Monitoring of the economic impact of tourism 6.3.b Encouraging innovation 6.4 Institutionalization of cooperation with the private sector 6.5.a Agility of the administrative management 6.5.b Adaptation of tourist rules/laws 6.6 Inter-administrative coordination	
7. Performance: social and economic outcomes	7.1 Total tourism revenues 7.2 Tourism model efficiency	7.1 Total tourism revenues 7.2.a Tourist revenues range of accommodation place 7.2.b Hotel profitability (RevPAR)	

Continued on next page

Table A.1 – continued from previous page

Pillars	Crucial areas for competitiveness	Indicators	Sub indicators
	7.3 Seasonality	7.3 Seasonality	
	7.4 Market positioning	7.4.a Average expenditure of tourists	
		7.4.b Tourist satisfaction	
	7.5 Social contribution	7.5.a Tourism revenues per inhabitant	
		7.5.b Tourism employment	

Source: Own elaboration

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 3, 2017, 175–200
DOI: 10.18335/region.v4i3.171

journal homepage: region.ersa.org

Spatial mismatch, wages and unemployment in metropolitan areas in Brazil

Ana Maria Bonomi Barufi[1], Eduardo Amaral Haddad[1]

[1] University of São Paulo, São Paulo, Brazil

Received: 25 October 2016/Accepted: 9 August 2017

Abstract. The spatial mismatch hypothesis states that a lack of connection to job opportunities may affect an individual's prospects in the labour market, especially for low-skilled workers. This phenomenon is especially observed in large urban areas, in which low-skilled minorities tend to live far away from jobs and face geographical barriers to finding and keeping jobs. This paper aims to investigate whether this negative relationship between spatial mismatch and labour market outcomes is valid in Brazil after controlling for individual characteristics. Our conclusions indicate that there is no clear relation between different measures of accessibility to jobs and the probability of being unemployed. However, for wages there is a clear correlation, which is stronger in larger metropolitan areas than in the country and has a more detrimental effect for low-skilled workers. This paper contributes to the literature by investigating the spatial mismatch in urban labour markets in Brazil. For the empirical literature in the country, this is an original contribution, as the comparison of intra-urban labour market dynamics of different urban areas provide a more comprehensive perspective of the role city size may play in local labour markets. Given the exploratory nature of this work, our results still rely on strong identification hypotheses to avoid potential bias related to simultaneous location decisions of workers and firms within the city. Even if these conditions do not hold, the results are still meaningful as they provide a better understanding of the conditional distribution of wages and the unemployment rate in the biggest metropolitan areas of Brazil.

JEL classification: R32, J64, J31

Key words: spatial mismatch, labour market, metropolitan areas.

1 Introduction

The spatial landscape of labour market opportunities varies significantly within an urban area. On average, the number of job openings and the wage level tend to decline as distance to the urban centre increases, which is usually modelled as a monocentric city. However, this relationship varies according to specific characteristics of each city, related to geography, amenities' distribution, sector composition and specialisation, transportation policies, the number of business centres (polycentric or monocentric city), among other factors (Capello 2007). Another important source of heterogeneity in the urban shape comes from the locational choices of firms in different sectors (McCann 2013), based on

their cost-benefit analysis coming from the interaction between land and transportation costs, and the potential benefits that may arise from a more central location.

In some cases, jobs with better pay in the service sector can be concentrated near the centre, as they benefit more from knowledge spillovers that generate agglomeration externalities (Partridge et al. 2009). On the other hand, manufacturers started moving to the outskirts of the bigger cities in order to avoid high rents, effect that is widely acknowledged in the literature, together with additional impacts on the housing market (Lucas Jr, Rossi-Hansberg 2002). Furthermore, this relationship is said to be stronger for larger and denser areas, because congestion costs and the size of the urban sprawl lead to a higher cost of living in central areas. In this context, the spatial mismatch relates the structure of cities to unemployment and poverty (Gobillon, Selod 2013).

The urbanisation process in Brazil was fast in the second half of the twentieth century, as the urbanisation rate went from less than 50% to more than 80% in forty years. More than 90% of the GDP is created in cities (Da Mata et al. 2007). However, this process was not accompanied by a similar rise in the country's GDP per capita (Chauvin et al. 2016). Urban areas with less than 100,000 inhabitants made up 23% of the Brazilian population in 2010, while in the US, they housed 33% of the population.

Local labour markets are formed by the interaction of firms and workers with heterogeneous skills in various geographical locations, given the strong connection between housing and labour markets. Geographical location gives market power to firms over potential workers, especially over those residing close to them. In their model, Brueckner et al. (2002) define two different spaces (skills spaces and urban spaces), and in equilibrium low-skilled workers will be distant from firms in both of these spaces, providing a rationale for socioeconomic ghettos (Zenou 2009), consistent with the spatial mismatch hypothesis (Kain 1968). The main underlying mechanism of this model is the monopsonistic power of firms in the surroundings close to them, which depends on the elasticity of the firm's labour pool (which itself is negatively related to the costs of commuting and acquiring skills). Brueckner et al. (2002) show that workers will be separated in space by skill type, and firms set wages that exploit this separation in space. Low-skilled workers will then live far away from their jobs.

There are at least two main dimensions through which this intra-urban equilibrium in the labour market can be evaluated: unemployment and wages. According to Zenou (1999), urban efficiency wages may lead to involuntary unemployment, as they are set above the competitive equilibrium wage in order to induce workers not to shirk. Moreover, individuals living far away from jobs have poor information about job opportunities, which decreases their probability of finding a job. As a result, spatial mismatch is observed in large urban areas in which low-skilled minorities live far away from jobs and face geographical barriers to finding and keeping jobs. In addition to the spatial dimension, there is also a social separation faced by low-skilled workers and minorities (Zenou 2013), which reduces their chances of finding a job.

Based on this theoretical perspective, this paper provides a two-fold analysis of the relationship between spatial mismatch and labour market outcomes in large metropolitan areas in Brazil. This effect is calculated through the relationship between the average wage or the probability of being unemployed and distance to jobs (measured as the commuting time from home to work or the distance to the main business centre). This paper therefore contributes to the literature by investigating the spatial mismatch in urban labour markets in Brazil. For the empirical literature in the country, this is an original contribution, as the comparison intra-urban labour market dynamics of different urban areas provide a more comprehensive perspective of the role city size may play in local labour markets.

Moreover, it shows empirically that in the Brazilian case the spatial mismatch is more relevant in relation to individual wages, while the probability of being unemployed is not as regularly distributed in space. The latter result is an interesting contribution to the literature and may indicate that the probability of unemployment may not be the best measure to attest the effect of the spatial mismatch in the labour market. According to the literature, duration of unemployment may be a better fit for this role, a variable that is not available in the database considered in this study.

There is also an emphasis on how the spatial mismatch can be more harmful to low-skilled workers, a result that is in accordance to previous findings in the literature. In sum, city size and the capacity that individuals have to adapt and find job opportunities are relevant aspects to be considered to understand intra-urban labour market dynamics

The paper is structured as follows. Section 2 provides a brief literature review of spatial mismatch and local labour markets focusing on social interactions within the city. In Section 3, we describe the econometric strategy and the database, while in Section 4 we analyse the results. Concluding remarks follow in Section 5.

2 Spatial mismatch and labour market equilibrium

The intra-urban spatial distribution of economic agents and production inputs has been modelled as the result of location decisions made by workers and firms (Roback 1982). A wide range of factors, among whose there are agglomeration economies, may be included in different models, as indicated by the New Economic Geography and Urban Economics literatures (Fujita, Thisse 2012, Krugman 1995, Ottaviano 2004). The locational problem is usually analysed by evaluating how local prices (rents and wages) relate to the distance from the present location to the Central Business District (CBD) of the city (Lucas Jr, Rossi-Hansberg 2002). Distance to multiple tiers of the urban hierarchy within a city can also be relevant for this analysis (Partridge et al. 2009).

The concept of spatial mismatch dates back to the mid-1960s (Kain 1992). This concept appears as a possible partial explanation to racial conflicts and riots in the United States, with the identification of ghettos and unequal labour market outcomes. Low rates of employment and low wages for Afro-American workers could be related to limitations on residential choice and the distribution of jobs around the city. Among other dimensions, education, housing and employment reflect and reinforce the spatial mismatch in cities. There has been significant discussion on whether this hypothesis does explain inequalities in the city, given the variety of analytical methods, spatial mismatch measures and data aggregation levels.

More than that, there was considerable uncertainty about the magnitude of the effects of spatial mismatch in urban areas (Holzer 1991). Recently, Kain (2004) showed that the public education system of the United States reinforced the spatial mismatch, given that racial segregation resulted in the concentration of Black children in low-achieving schools. More recent developments combine the concept of spatial mismatch with the analysis of local prices within a city and the embedded location decisions of workers and firms. Spatial mismatch in the labour market means that people face spatial frictions when accessing jobs in metropolitan areas (Houston 2005a). This phenomenon relates to the way in which low-skilled minorities are affected by distance to job locations (Zenou 2009). The resulting distributions arise from the equilibria in the labour and the housing markets, which are simultaneously determined by the different decisions made by firms and workers.

The spatial mismatch hypothesis argues that low-skilled minorities face poor labour market outcomes because they are disconnected from job opportunities within the city (Gobillon et al. 2007). Even nowadays, this concept is still commonly used to investigate the case of afro-descendent population or other minorities in US cities, who often live far away from low-skilled jobs that are available in the suburbs of American cities (see for instance Ihlanfeldt 2006, Zenou 2009, Andersson et al. 2014).

The range of mechanisms underlying the theoretical frameworks that generate spatial mismatches are related either to the labour market itself or to the factors that potentially explain why minorities are physically disconnected from jobs (Gobillon, Selod 2013). According to Gobillon et al. (2007), these mechanisms can be analysed separately for workers and firms. From the workers' perspective, they are the following:

(i) long commuting may lead a worker to refuse a job opportunity after carrying out a cost-benefit analysis;

(ii) search efficiency may decrease with distance to jobs;

(iii) search intensity may also be affected by distance to jobs; and

(iv) high search costs may cause workers to restrict their search to a limited area.

From the firms' perspective, the main mechanisms are:

(v) stigma or prejudice may make firms discriminate against workers who live in certain locations;

(vi) employers may pay lower wages or refuse to hire workers who commute for long distances, as the commuting may decrease their productivity; and

(vii) employers may have a prejudice against specific workers because of the expected preferences of their customers.

As mentioned above, the spatial mismatch hypothesis is usually considered in the specific case of low-skilled minorities living in urban areas. However, the concept of 'spatial mismatch' in general terms is broadly used to investigate the disparity between locations of jobs and individuals that lead in an endogenous way to different levels of unemployment and wages across a city.

Among some of the theoretical models devoted to describing spatial mismatches in the urban environment, Zenou (2000) develops a model with endogenous city formation mechanisms that result in jobs concentrating in the CBD, employed individuals residing in the vicinity of the city centre, and the unemployed being further away from jobs. Urban unemployment will then be reinforced in the outskirts of the city, because the further away an individual is from jobs (which are concentrated in the CBD), the harder it is for her to find a job. Within a similar setting generated from a model based on a monocentric city combined with an efficiency wage mechanism and high reallocation costs, wages are expected to decrease with distance to the centre, as demonstrated by Zenou (2006).

It is important to note that the modelling of metropolitan labour markets can be significantly different for low-skilled and high-skilled workers, given the more limited distance that low-income individuals can commute. Thus, low-skilled workers will face a segmented urban labour market, while for high-skilled workers space is less restrictive. Unemployment for low-skilled workers will be associated with the lack of jobs in the areas close to their residence, while high-skilled workers will search for jobs in a wider spatial scale (Morrison 2005). Therefore, for high-skilled individuals, urban landscape is expected to have a smaller impact on their labour market outcomes. These two mechanisms can co-exist within the city to generate the observed distribution of unemployment rates.

One additional remark is that the literature of spatial mismatch is intrinsically related to spatial spillovers, social networks and proximity in different dimensions (Topa, Zenou 2015). Accessibility to jobs captures these effects just in a partial way, as it differentiates individuals by their reach to opportunities, and an interesting extension of research in this area should encompass these neighbouring relations in a more direct way.

Despite the large amount of empirical literature, Houston (2005b) argues that there is no clear consensus on the importance of the spatial mismatch in the explanation of labour market outcomes. Andersson et al. (2014) consider the duration of unemployment as a labour market outcome to measure the effects of spatial mismatch. They use a matched employer–employee database, and build person-specific measures of job accessibility with an empirical model of transport modal choice and network travel-time, finding that better job accessibility helps to decrease the duration of joblessness for lower-paid workers. Moreover, under-privileged groups are more affected by the lack of accessibility. The same dependent variable is employed by Rogers (1997), whose results indicate that unemployment duration in the Pittsburgh labour market area is influenced by residential location relative to employment opportunities, especially for less-educated individuals. According to Johnson (2006), the efficiency of job search is largely related to job accessibility. Then, 40% of the racial disparities in search duration is explained by spatial search-related variables.

The total number of jobs available in each region of the city and the impedance for reaching those regions can be used to define accessibility to jobs in a specific location. The impedance measure is usually defined either by the Euclidean distance or by commuting time between residential location and jobs, which may be derived from transport

availability in each area of the city. The latter approach is followed by Vieira, Haddad (2015) for the São Paulo Metropolitan area, and they find indications that accessibility and income are strongly and positively related in the city. Di Paolo et al. (2016) find that car availability is relevant for job–education mismatch and that public transportation has an effect on better matching in the labour market for each schooling level.

Åslund et al. (2010) calculate the accessibility measure by considering the number of jobs and people of working age within a 5 kilometres-radius of the individual's residential location. They consider the exogenous allocation of refugees in Sweden ten years earlier and build an instrument that is based on how accessible jobs are to immigrants in their arrival year, and find a positive correlation between local job proximity and individual outcomes.

Job accessibility, demand and supply in the Chicago metropolitan area are used by Hu (2014) to find that socioeconomic restructuring (an increase in poverty and a reduction in relevant job opportunities) negatively affects poor job seekers, while spatial transformation (when jobs and job seekers move to the outskirts of the city) has a positive effect on their job prospects. The latter effect is caused by poorer individuals following jobs to suburban areas. With a similar empirical strategy, Hu, Giuliano (2014)'s results indicate that there is no relationship between spatial accessibility and the unequal employment status of the poor in the Los Angeles metropolitan area.

According to Tyndall (2015), public transportation has a causal and negative effect on neighbourhood unemployment rates, particularly for groups who are more dependent on this transport mode. The author explores a natural experiment from Hurricane Sandy, which exogenously reduced access to public transport in some neighbourhoods in New York City.

The empirical literature on spatial mismatch can be subdivided into two main strands: the first aims to understand the causes, while the second discusses the consequences of a spatial mismatch (Gobillon, Selod 2013). Houston (2005b) states that the consequences of a spatial mismatch are usually evaluated through an analysis of (i) residential segregation, (ii) comparisons of commuting times, (iii) comparisons of earnings, and (iv) measures of job proximity. Accordingly, Ihlanfeldt (2006) highlights the fact that the effects of spatial mismatch have been investigated on lower earnings, longer commutes and higher unemployment, especially in the case of black workers in the United States. Usually, employment and earnings equations include measures of local job opportunities, with a strategy based on a gravity model with a distance-decay function to take account of being further away from job opportunities.

Among the main econometric problems arising from this strategy there is the fact that residential location and the measurement of job opportunities are potentially endogenous (Ihlanfeldt 2006). Such endogeneity may appear through self-selection of more or less productive workers to specific areas, by the potential reverse causality of job opportunities and the probability of being unemployed, or through the simultaneous location decisions of firms and workers in a general equilibrium setting. One can deal with the simultaneity issue by including historical or geographical instruments that influenced the location of transportation infrastructure within a city without directly determining the location of workers and firms. This approach is explored by Haddad, Barufi (2017) for São Paulo Metropolitan Region with river shore access as an instrument, but is not replicable for the whole country as such detailed geographical information is not available yet in a larger scale.

Our identification strategy will be based in more restrictive hypotheses. In the short run, prices in the labour market are assumed to be close to the equilibrium level, and workers and firms are relatively immobile (Gibb et al. 2014). This endogeneity issue is then expected to be less relevant in the case of labour market outcomes. In addition, the measurement of local job opportunities can be indirect (using the assumption that there is a geographical centre in the city or by considering commuting time as a possible measure of the distance to jobs). The specific location of job opportunities is not included in the analysis, meaning that this endogeneity issue can be less relevant. In this study, we will assume that these aspects are able to soften such concerns. In any case, the potential direction of an endogeneity bias will be discussed in the following sections.

Usual measures of spatial mismatch may be problematic (Houston 2005b). On the one hand, long commutes may be a sign of either high mobility (highly paid workers) or a spatial mismatch between workers and jobs. On the other hand, different groups have specific propensities to commute, which means that studies usually measure commuting patterns of employed individuals, while spatial mismatch is generally concerned with the unemployed, who may behave differently. Houston (2005b) also suggests that job accessibility should take into account not only distance but also the amount of competition for the accessible jobs. Finally, total travel burden should take into account time, pecuniary costs and inconvenience (Bruzelius 1979). Commuting time, cost or distance are therefore, by themselves, incomplete measures. However, data availability restricts this analysis to such incomplete measures. We acknowledge this limitation and try to assess its potential impact in our results.

In summary, the empirical literature finds some mixed results, especially regarding the relationship between different measures of spatial mismatch and the unemployment rate. However, an increase in accessibility to jobs seems to improve labour market outcomes, especially for low-skilled minorities for whom the spatial mismatch is more relevant. There are significant empirical issues related to the estimation of this effect, whose consequences will be further discussed.

The next section presents our empirical strategy, which deals with comparisons of earnings and measures of job proximity (items (iii) and (iv) discussed above and listed by Houston (2005b)). In addition, we focus on the probability, for each economically active individual, of being unemployed, according to her residential location. To compare earnings, the unavailability of data means that we measure wages from a residential location perspective instead of a workplace basis, even if the latter would be a more appropriate approach (Houston 2005b).

3 Empirical strategy and data

The empirical strategy developed here is based on the estimation of the relationship between different measures of distance to jobs and labour market outcomes (earnings and the probability of being unemployed). All dependent variables are residence-based, due to data availability. Such strategy aims at exploring different dimensions of the spatial mismatch hypothesis in Brazilian metropolitan areas.

Estimations are conducted for individuals residing in a specific metropolitan area in order to capture the effect of each variable in relative terms within a specific urban structure. We assume that the wage equation can be written as follows:

$$w_i = \alpha + \beta X_i + \gamma_1 \text{inv\_dist}_r + \gamma_2 \text{inv\_dist}_r^2 + \epsilon_i \tag{1}$$

where w_i is the logarithm of the hourly wage measured for employed individuals who do not work at home, and X_i includes age, age squared, sector of activity, occupation, formalization status of the job, colour or race, education level, whether the individual is married, whether he or she has at least one child younger than fifteen living in the house, whether the house is owned by the family and whether the person is or is not the head of the household. In addition, inv_dist_r refers to the inverse of the Euclidean distance from the centroid of the weighting area to the main business centre[1]. This strategy is adopted since there is no data available to measure distance over each city's road infrastructure.

An alternative formulation for the reduced form presented in (1) is given by:

$$\begin{aligned} w_i \quad = \quad & \alpha + \beta X_i + \theta_1 \text{time\_commut\_6\_30}_i + \theta_2 \text{time\_commut\_31\_60}_i \\ & + \theta_3 \text{time\_commut\_61\_120}_i + \theta_4 \text{time\_commut\_121\_more}_i + \epsilon_i \end{aligned} \tag{2}$$

In this case, instead of the inverse distance to the centre, commuting time from home to work is used to evaluate the relationship between wages/productivity and the urban

[1] Under the simplifying assumption of a monocentric city, we consider the inverse distance from the weighting area where the individual lives to the main business centre of the metropolitan area, to calculate an approximate measure of distance to jobs.

landscape[2]. All these models are estimated with a simple OLS. Another dimension of spatial mismatch is the heterogeneity in the unemployment rates within the urban area. This dimension will be assessed by estimating the probability of being unemployed for each economically active individual, given her relative location to the main centre of the city:

$$h_i = P[U_i = 1] = F[\beta X_i + \gamma_1 \text{inv\_dist}_r + \gamma_2 \text{inv\_dist}_r^2] \quad (3)$$

In this specification, U_i refers to the employment status (it equals 1 when a person is unemployed) and F is a logistic cumulative probability function. Here, X_i is the set of observed characteristics for the individual (age, age squared, colour or race, education level, whether the individual is married, whether he or she has at least one child younger than fifteen living in the house, whether the house is owned by the family and whether the person is or is not the head of the household). Finally, β is a vector of parameters, and inv_dist_r is measured as before. An alternative formulation is the following:

$$\begin{aligned} h_i &= P[U_i = 1] \\ &= F[\beta X_i + \theta_1 \%\text{\_time\_commut\_6\_30}_r + \theta_2 \%\text{\_time\_commut\_31\_60}_r + \\ &\quad \theta_3 \%\text{\_time\_commut\_61\_more}_r] \end{aligned} \quad (4)$$

In this case, the spatial mismatch is approximated by the percentage of individuals in the neighbourhood whose time spent in commuting belongs to a particular time span.

In sum, two different measures of accessibility to jobs are considered here. Individuals in the Demographic Census are located in weighting areas, as it will be better explained below. Then, the first measure is based on the Euclidean distance of the centroid of the weighting area of residence to the business centre of each metropolitan area. This centre is equivalent to the geographic coordinates of the administrative centre of the municipality with the largest employment level of each metropolitan area.

The second accessibility measure is calculated through the commuting time spent from home to work. As a limitation, this variable is only available in categories (up to five minutes, from six to thirty minutes, thirty minutes to one hour, more than one hour to two hours, more than two hours). In the case of wage models, this variable is obtained through the individual's own reported commuting time. For the probability of unemployment, it is calculated by the percentage of workers who reside in each weighting area that are classified in each category and used in the regressions for the individuals living in that specific weighting area.

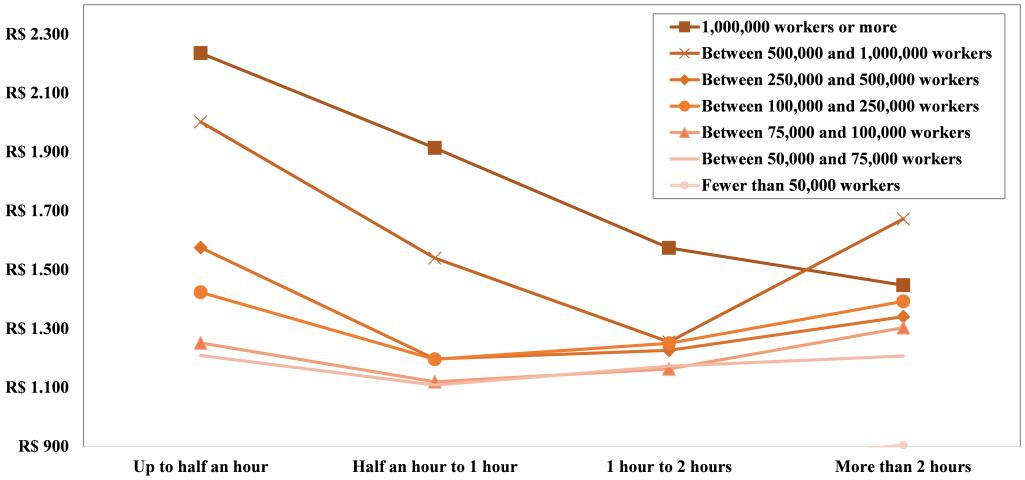
Apart from the whole database, these four models will be estimated for each metropolitan area and for three separate groups: (i) individuals who did not complete primary school[3], (ii) up to high school graduates without a college degree, and (iii) individuals who completed college education. In a country such as Brazil, inequality derived from the spatial mismatch can be more or less pronounced depending on the city size and the distance to the main concentration of job opportunities, and it may affect distinct skilled groups in different ways.

3.1 Database

The Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística – IBGE) conducts a Demographic Census every ten years, with regional disaggregation at the municipal level (or at the census area level for bigger municipalities). The Demographic Census collects information on the main characteristics of individuals and households, providing details on the living conditions of the population in each municipality, and serving as a very important policy instrument in a country with a land area the size of Brazil. A shorter questionnaire applies to the whole population at the

[2]This impedance measure is the commuting time from home to work, calculated at the individual level for the wage equation or for the neighbourhood in the case of the estimation of unemployment probability. This second approach may be associated with a multicentric city structure. Census data only provides commuting time in categories, what represents an additional limitation of this analysis.

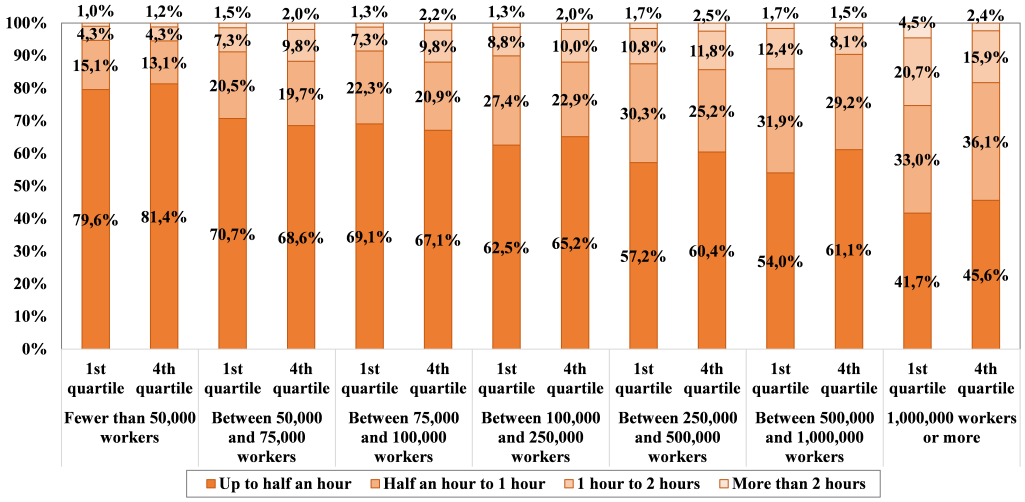
[3]Eight years of education.



Source: IBGE

Figure 1: Average wage of workers according to their commuting time from home to work and the size of the municipality of residence, 2010

census tract level, while specific individual characteristics are investigated in a longer set of questions that are given to a sample and are representative at the weighting areas level (conglomerates of census tracts with at least 400 households). Microdata at the individual level are available for this sample. We will use weighting areas as our definition of neighbourhood.

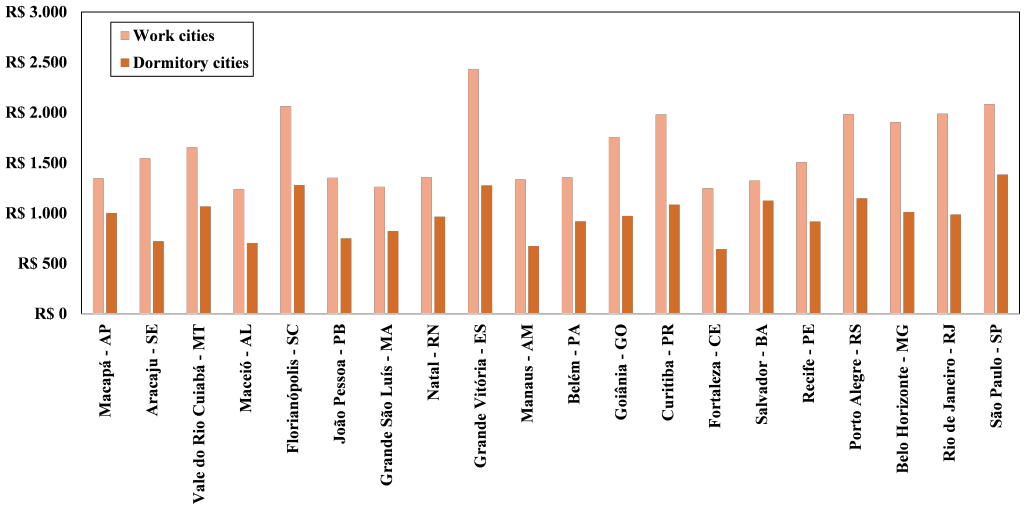


Source: IBGE

Figure 2: Distribution of workers who commute from home to work and belong to the 1st and the 4th quartile of the wage distribution, according to their commuting time and the size of the workforce in the municipality of residence, 2010

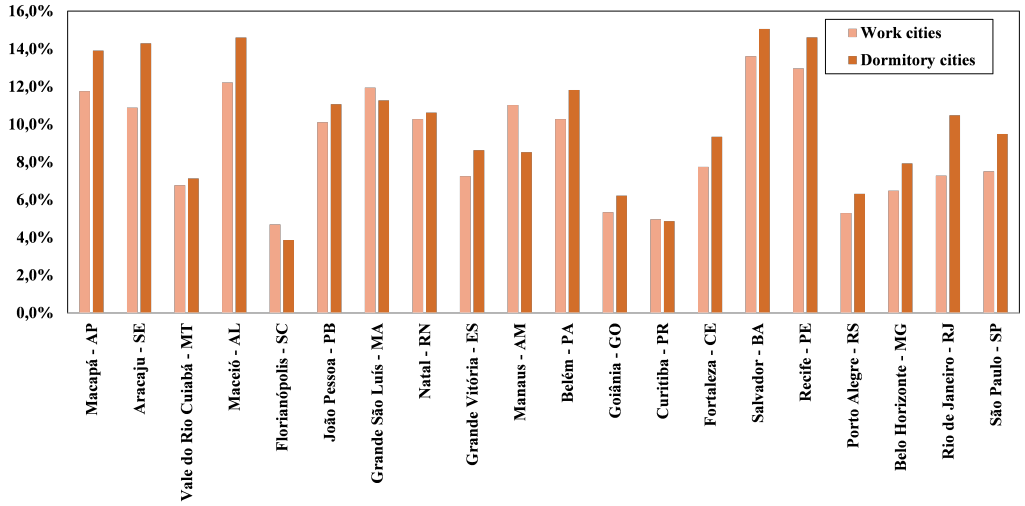
3.2 Descriptive statistics

The problem at hand is fundamentally related to metropolitan areas, as commuting costs and agglomeration economies become more relevant at a larger urban scale (Partridge et al. 2009). In fact, if one considers the average wage received by workers according to their commuting time from home to work, it is noticeable that the negative relationship between these two variables is clearer when cities with at least 500,000 workers are taken into account (Figure 1).



Source: IBGE

Figure 3: Average monthly wage for workers who live in work or dormitory cities inside each metropolitan area, (ordered by the size of working population), 2010



Source: IBGE

Figure 4: Average unemployment rate for people who live in work or dormitory cities inside each metropolitan area, (ordered by the size of working population), 2010

This difference between cities of different sizes is made clear in the analysis presented in Figure 2. In fact, the biggest differences in commuting times faced by workers in the richest (4^{th}) and the poorest (1^{st}) quartiles of the wage distribution in each municipality is seen in places with at least 500,000 workers. Furthermore, the decreasing relationship between wages and commuting time is stronger for those who commute for up to two hours.

For this reason, only 20 metropolitan areas containing state capitals were included in the study. In addition, only male workers aged 25 to 64 years old were kept in the database, in order to homogenise their decisions to participate in the labour market. For the wage regression, the database contained only workers who commuted to work and returned home every day.

It is also possible to show how wages and the unemployment rate vary according to the distance between the residential location of a worker and the centre of the city. Considering the daily commuting flows from home to work obtained from the Demographic

Table 1: Descriptive characteristics of each metropolitan area (ordered by the size of working age population), 2010

Metropolitan region	Macro region	Average hourly wage (R$ 2010)	Unemployment rate	Individuals commuting >1 hour (in %)	Working age population (men aged 25-64)
Macapá - AP	North	R$ 10.44	7.7%	5.3%	85,494
Aracaju - SE	Northeast	R$ 10.87	7.4%	10.7%	159,838
Vale do Rio Cuiabá - MT	Centre-West	R$ 13.58	4.3%	7.7%	160,638
Maceió - AL	Northeast	R$ 9.27	8.2%	13.3%	216,904
Florianópolis - SC	South	R$ 13.77	2.6%	6.6%	217,208
João Pessoa - PB	Northeast	R$ 9.72	6.5%	7.7%	230,930
Grande São Luís - MA	Northeast	R$ 10.96	7.5%	16.1%	244,017
Natal - RN	Northeast	R$ 9.85	7.1%	8.4%	258,207
Grande Vitória - ES	Southeast	R$ 11.94	4.9%	14.6%	353,561
Manaus - AM	North	R$ 11.19	7.1%	16.7%	378,496
Belém - PA	North	R$ 10.85	7.0%	14.4%	402,170
Goiânia - GO	Centre-West	R$ 12.32	3.4%	11.2%	415,541
Curitiba - PR	South	R$ 13.51	3.0%	13.1%	623,103
Fortaleza - CE	Northeast	R$ 9.41	5.6%	12.4%	666,504
Salvador - BA	Northeast	R$ 11.01	9.2%	20.0%	723,297
Recife - PE	Northeast	R$ 10.00	9.5%	17.2%	745,952
Porto Alegre - RS	South	R$ 12.38	3.7%	11.4%	807,268
Belo Horizonte - MG	Southeast	R$ 11.82	4.2%	18.7%	1,115,715
Rio de Janeiro - RJ	Southeast	R$ 12.92	5.8%	30.5%	2,402,075
São Paulo - SP	Southeast	R$ 15.37	5.7%	28.8%	3,953,270

Source: IBGE

Census of 2010, it is possible to define work and dormitory cities in each metropolitan area. The former are characterized by a higher inflow of people going there to work than an outflow of those who live there and go somewhere else to work, while the latter present a higher daily worker outflow than an inflow.

Figure 3 shows that average wages are much higher for people who live in work cities than for those who live in dormitory cities. However, in the case of the unemployment rate, there are mixed signs (Figure 4). In some metropolitan areas (Manaus, Grande São Luís, Florianópolis and Curitiba), dormitory cities show a lower unemployment rate than work cities. This pattern is unexpected under the hypothesis of a monocentric metropolitan area, but may be associated to the fact that these specific metropolitan areas are less dense than other more developed metropolitan areas in Brazil, for which the unemployment rate is larger in dormitory cities.

The econometric discussion outlined above explains the need to calculate the distance of each weighting area to the relevant business centre. This should be done on the basis of the main location of jobs around the city. In Brazil, however, there is no consolidated database covering all metropolitan areas and showing the location of jobs. Therefore, we consider a different approach, in which the centre of the metropolitan area is given by the administrative centre of the largest municipality (defined according to the number of employed individuals in 2010[4]).

Focusing more specifically on the models, the main descriptive characteristics are presented in Tables 1, 2 and 3, and Tables A.1, A.2 and A.3 in the Appendix. Table 1 indicates that the metropolitan areas considered in this study are significantly heterogeneous and should be treated separately, as each of them has a specific distribution of jobs and wages. Furthermore, areas with a bigger labour market have a higher average wage and a higher percentage of workers who commute for more than one hour to reach their jobs. This characteristic is clearer for metropolitan areas with more than a million male workers aged 25 to 64. For the unemployment rate, there seems to be more of a

[4]Data obtained from the Ministry of Labour and available at http://pdet.mte.gov.br/acesso-online-as-bases-de-dados.

Table 2: Percentage of workers who spend more than one hour commuting from home to work, according to the distance the worker lives from the centre, (ordered by the size of working population), 2010

	Distance from centre (in kilometer)							
	<2.5	2.5 to <5	5 to <10	10 to <20	20 to <30	30 to <40	40 to <50	50 or more
Macapá - AP	4.4%	4.9%	5.5%	6.3%		8.1%		4.0%
Aracaju - SE	7.0%	8.1%	11.4%	15.3%	15.9%			
Vale do Rio Cuiabá - MT	6.0%	5.0%	7.2%	10.2%	15.8%			8.7%
Maceió - AL	5.6%	6.9%	10.2%	20.4%	14.2%	11.9%		
Florianópolis - SC	3.5%	3.0%	5.2%	9.4%	12.5%	4.1%	4.9%	2.5%
João Pessoa - PB	6.7%	6.8%	8.0%	8.7%	6.0%	7.0%	10.7%	7.7%
Grande São Luís - MA	6.0%	11.2%	12.0%	23.2%	21.4%	14.5%		
Natal - RN	16.3%	13.4%	5.8%	5.4%	9.6%	8.3%	6.2%	
Grande Vitória - ES	7.4%	9.9%	16.9%	16.3%	22.2%	11.4%	6.1%	
Manaus - AM	10.6%	9.4%	12.0%	22.7%			20.0%	11.4%
Belém - PA	5.8%	6.7%	7.8%	19.4%	23.0%	11.2%	15.2%	
Goiânia - GO	3.6%	4.5%	6.6%	15.7%	21.2%	12.2%	11.5%	8.2%
Curitiba - PR	3.8%	4.2%	9.8%	16.5%	21.1%	6.9%	17.7%	8.7%
Fortaleza - CE	14.7%	13.2%	14.1%	12.6%	8.7%	8.0%	6.4%	5.3%
Salvador - BA	15.2%	13.8%	19.8%	29.4%	18.8%	9.0%	9.4%	10.7%
Recife - PE	7.4%	8.1%	13.1%	24.1%	20.1%	15.9%	7.2%	12.6%
Porto Alegre - RS	2.5%	5.1%	8.4%	15.7%	17.6%	7.8%	4.3%	4.7%
Belo Horizonte - MG	7.3%	10.9%	13.0%	24.6%	25.8%	17.5%	13.4%	6.3%
Rio de Janeiro - RJ	14.1%	13.5%	18.3%	28.1%	37.8%	39.7%	39.4%	22.8%
São Paulo - SP	23.9%	20.0%	21.1%	28.5%	34.7%	29.9%	21.5%	16.9%

Source: IBGE

regional aspect to the level observed in each metropolitan area, as regions located in the Northeast, for example, show a much higher level of unemployment than other regions.

There is a strong relationship between commuting time and distance to the centre, as can be seen in Table 2. In São Paulo and Rio de Janeiro, the largest metropolitan areas in Brazil, the percentage of individuals who commute for more than one hour is significantly higher for people who live more than 10km from the centre than for those living less than this distance away. However, this percentage decreases when the distance to the centre is greater than 30km in São Paulo or 40km in Rio de Janeiro. Since our objective is to investigate labour market characteristics related to the main business centre of each metropolitan area, we will focus on individuals living within a circle with a radius of 30km.

Table 3: Descriptive statistics by individual characteristics, 2010

	Unemp. Rate	Average hourly wage (R$ 2010)
Age		
25 to 34 years old	7.4%	9.75
35 to 44 years old	4.9%	12.31
45 to 54 years old	4.7%	15.48
55 to 64 years old	4.7%	19.06
Education level		
Less than 7 years of schooling	6.9%	6.62
8 to 10 years of schooling	6.3%	8.29
11 to 14 years of schooling	5.7%	11.21
15 years of schooling or more	3.0%	33.37

Continued on next page

Table 3 – continued from previous page

	Unemp. Rate	Average hourly wage (R$ 2010)
Colour		
White	4.8%	16.91
Black	6.8%	8.43
Yellow	5.1%	18.20
Brown	6.7%	8.84
Indigenous	6.4%	8.97
Marital status		
Single	7.7%	10.08
Married	3.7%	15.48
Children		
No children up to 15 years old	6.9%	13.07
Has at least one child up to 15 years old	4.1%	12.29
Home ownership		
Tenant	5.4%	11.39
Owned home	5.9%	13.21
Household position		
Another member of the household	8.0%	10.36
Head of the household	4.2%	14.33
Formality status		
Informal sector		9.46
Formal sector		13.93
Sector		
Agriculture		7.51
Manufacture and construction		9.59
Other industrial activities		14.26
Commerce		10.26
Services		10.53
Auxiliary services		17.83
Transport and communication		9.86
Health and social services		24.84
Education		17.85
Public sector		22.01
Other activities		15.79
Occupation		
Non-applicable		16.80
Leaders		30.45
Scientific, artistic or similar		30.88
Technical level		14.93
Administrative service		9.61
Commerce and service		7.14
Agriculture, livestock, extractive activities		4.14
Manufacture		7.26
Military		23.73
Commuting time to work		
Up to 5 minutes		13.66
6 to 30 minutes		13.47
31 minutes to 1 hour		12.68
More than 1 hour to 2 houres		11.27
More than 2 hours		11.15

Source: IBGE
Notes: The unemployment rate is calculated for the weighting area in which the individual resides

In Table 3, we can note that the wage level is higher for older individuals, those who are better educated, married people, those who are Indians, from Asiatic ancestry or white, those who are the head of a household, people employed in the formal sector and those who work in health and social services or leaders, scientists or artists. In addition, workers who commute for a longer time have a lower salary, on average. On the other hand, the unemployment rate is higher for younger individuals, those who are less educated, those who are black or brown, single people, people with no children, and those who are not heads of households.

The theory of spatial mismatch states that a lack of connection to job opportunities may affect an individual's prospects in the labour market, especially for low-skilled workers. Complementing the results presented in Table 3, Tables A.1, A.2 and A.3 provide wage levels and unemployment rates using different impedance measures. Distance to jobs can be calculated in many ways: (i) distance from the centroid of the weighting area to the business centre of the metropolitan area; (ii) individual commuting time from home to work; or (iii) percentage of workers in the weighting area whose commuting time falls within each time span. For the wage equation, we consider alternatively (i) and (ii) for employed individuals. On the other side, for the estimation of the probability of unemployment, (i) and (iii) are used, calculated at the weighting area level.

With these considerations in mind, Tables A.1, A.2 and A.3 show that wages seem to be higher near the centre of each metropolitan area, and that this effect is stronger in larger areas. However, for the unemployment rate, the expected positive relationship with distance to jobs is not clear. The main results will be presented in the next section.

4 Results

The first set of results refers to the estimation of wage equations that control for individual characteristics and uses two different measures of relative distance in the city: the distance to a unique centre (a monocentric city) and the distance to each worker's job (a multicentric city).

Table 4 shows that wages have a positive relationship with the inverse distance to the main centre of each metropolitan area (and, as a consequence, a negative relationship with distance itself). This effect is more significant for larger metropolitan areas, and it seems to be stronger for individuals with a higher education level. Therefore, wages are lower for individuals who live further away from the main business centre. However, this result demonstrates more of a correlation than a causal effect, especially because individuals are analysed with reference to their residential location. There may be inverse causality in this case, as an individual's choice of location may be affected by the wage previously received, and this can affect current labour market prospects and productivity.

Table 4: OLS regressions of the logarithm of the hourly wage, for all individuals and by education group

	Macapá - AP	Aracaju - SE	Vale do Rio Cuiabá - MT	Maceió - AL	Florianó-polis - SC	João Pessoa - PB	Grande São Luís - MA
All individuals							
Inverse of distance	0.244‡	0.935‡	1.323‡	0.439‡	0.281‡	0.004	0.122
Inverse of distance squared	-0.072	-0.855‡	-2.057‡	-0.064	-0.053	-0.004	-0.007
N	5,559	7,736	8,121	9,068	15,481	10,490	10,421
Adjusted R squared	0.429	0.463	0.364	0.455	0.421	0.44	0.354
Up to incomplete primary school							
Inverse of distance	0.349†	1.482‡	0.627	0.096	-0.308†	0.428‡	0.546‡
Inverse of distance squared	-0.126	-2.011‡	-0.878	0.234	0.213*	-0.118‡	-0.030‡
N	1,754	2,889	2,777	3,918	4,158	4,551	2,916
Adjusted R squared	0.134	0.121	0.088	0.103	0.081	0.116	0.081
Complete primary school to incomplete tertiary school							
Inverse of distance	0.215*	0.534*	0.902†	0.568‡	0.151*	0.008	-0.046
Inverse of distance squared	-0.058	-0.199	-1.604†	-0.244*	-0.033	-0.001	0.002

Continued on next page

Table 4 – continued from previous page

N	3,015	3,979	4,187	4,091	8,071	4,674	6,403
Adjusted R squared	0.321	0.32	0.2	0.276	0.23	0.243	0.197
Complete tertiary school							
Inverse of distance	0.246	1.458†	3.897‡	0.740‡	0.725‡	-1.237‡	-0.152
Inverse of distance squared	-0.083	-1.601†	-5.515‡	-0.257	-0.274‡	0.316‡	0.007
N	790	868	1,157	1,059	3,252	1,265	1,102
Adjusted R squared	0.293	0.283	0.196	0.296	0.3	0.298	0.237

	Natal - RN	Grande Vitória - ES	Manaus - AM	Belém - PA	Goiânia - GO	Curitiba - PR	Fortaleza - CE
All individuals							
Inverse of distance	-0.179	0.300‡	0.633‡	0.594‡	1.283‡	1.069‡	-0.355‡
Inverse of distance squared	-0.211	-0.344‡	-0.347‡	-0.258‡	-0.739‡	-0.302‡	0.224*
N	12,056	24,887	11,912	15,523	17,317	32,745	26,254
Adjusted R squared	0.46	0.431	0.34	0.374	0.361	0.389	0.408
Up to incomplete primary school							
Inverse of distance	0.365*	-0.143	1.001‡	0.461‡	1.252‡	0.764‡	-0.341†
Inverse of distance squared	-0.714†	0.253	-0.492‡	-0.214*	-0.812‡	-0.174‡	0.364*
N	4,512	7,780	3,541	4,933	6,466	10,605	9,248
Adjusted R squared	0.124	0.094	0.097	0.077	0.071	0.092	0.076
Complete primary school to incomplete tertiary school							
Inverse of distance	-0.613‡	0.202†	0.471‡	0.439‡	1.175‡	1.056‡	-0.023
Inverse of distance squared	0.278	-0.290†	-0.238‡	-0.148	-0.674‡	-0.293‡	-0.161
N	6,041	13,138	6,989	8,728	8,448	16,835	14,345
Adjusted R squared	0.268	0.216	0.182	0.19	0.215	0.19	0.213
Complete tertiary school							
Inverse of distance	0.334	1.460‡	0.445	1.134‡	1.255‡	1.038‡	-1.937‡
Inverse of distance squared	-1.546	-1.576‡	-0.333*	-0.646‡	-0.716‡	-0.295‡	1.361†
N	1,503	3,969	1,382	1,862	2,403	5,305	2,661
Adjusted R squared	0.268	0.283	0.216	0.256	0.249	0.254	0.269

	Recife - PE	Salvador - BA	Porto Alegre - RS	Belo Horizonte - MG	Rio de Janeiro - RJ	São Paulo - SP
All individuals						
Inverse of distance	0.758‡	0.538‡	1.251‡	1.267‡	0.617‡	1.399‡
Inverse of distance squared	-0.639‡	-0.202‡	-0.616‡	-0.681‡	-1.005‡	-1.107‡
N	33,635	25,865	35,715	47,034	79,277	154,088
Adjusted R squared	0.409	0.416	0.435	0.427	0.404	0.369
Up to incomplete primary school						
Inverse of distance	0.07	0.177	0.411*	0.362†	-0.013	1.359‡
Inverse of distance squared	-0.001	0.12	-0.118	0.045	0.036	-1.011‡
N	11,393	7,617	10,425	17,149	20,809	45,808
Adjusted R squared	0.093	0.088	0.091	0.079	0.073	0.08
Complete primary school to incomplete tertiary school						
Inverse of distance	0.607‡	0.379‡	1.234‡	1.101‡	1.066‡	1.650‡
Inverse of distance squared	-0.549‡	-0.112†	-0.575‡	-0.600‡	-1.686‡	-1.271‡
N	18,298	14,563	20,476	23,695	43,703	79,875
Adjusted R squared	0.203	0.217	0.233	0.216	0.186	0.18
Complete tertiary school						
Inverse of distance	1.776‡	1.202‡	1.392‡	1.821‡	1.165‡	0.705‡
Inverse of distance squared	-1.734‡	-0.552‡	-0.717‡	-1.079‡	-4.248‡	-0.641‡
N	3,944	3,685	4,814	6,190	14,765	28,405
Adjusted R squared	0.254	0.246	0.244	0.291	0.228	0.202

Source: Authors' calculations

Notes: Controls: age, age squared, colour or race, household head, with children up to 15 years old, married, sector of activity, occupation, existence of a formal contract. For the regressions with all individuals, the education attainment of the individual was included as an additional control. Significance levels: * p<0.10, †p<0.05, ‡p<0.01. Only male individuals aged 25 to 64 years old living within a distance of 30km from the centre are considered in the analysis. Sampling weights are taken into account with Stata command pweight. Complete tables can be requested from the authors.

This issue may also be present when the spatial mismatch is captured by each individual's commuting time from home to work (Table 5). The estimated coefficients are then likely to be underestimating the real effect. Therefore, if this reverse causality issue is correctly dealt with, distance to jobs should be even more relevant in determining wage levels, as it would be possible to discount the effect of relocation by looking at job opportunities over the city.

In any case, Table 5 shows that the negative effect of commuting time on wages is significant for workers commuting for 30 minutes or more, and is higher the longer the time spent in this activity. For low-skilled workers in smaller metropolitan areas, wages are not significantly correlated to this measure of spatial mismatch. Moreover, for most metropolitan areas, workers who commute for two hours or more do not see any significant effect on their wages, which may result from the fact that there are only a few workers belonging to this group, and no clear wage pattern.

The second set of results refers to the probability of being unemployed. Coefficients are presented as odds-ratios, with values greater than one indicating a positive effect of the variable of interest on the probability of unemployment. Tables 6 and 7 present the estimated coefficients related to specific distance measures. Metropolitan areas are ranked from left to right according to the size of their labour market. There is an indication in Table 6 that the probability of unemployment is not significantly correlated with the inverse distance to the centre. This result is consistent for most metropolitan areas, and there is no specific pattern for groups with different levels of schooling. The same result is found when distance to jobs is measured by the time spent by workers in the neighbourhood commuting from home to work (Table 7). Once again, for most metropolitan areas this relationship is not significant, and it does not show any pattern regarding education level, labour market size, or the sign of the correlation itself in cases when it is in fact significant.

Table 5: OLS regressions of the logarithm of the hourly wage, for all individuals and by education group

	Macapá - AP	Aracaju - SE	Vale do Rio Cuiabá - MT	Maceió - AL	Florianópolis - SC	João Pessoa - PB	Grande São Luís - MA
All individuals							
Workers commuting 6'-30'	-0.093‡	0.002	-0.038	-0.029	-0.008	0.021	-0.013
Workers commuting >30'-1 hour	-0.100†	-0.058	-0.137‡	-0.017	-0.040*	-0.015	-0.022
Workers commuting >1-2 hours	-0.194‡	-0.069	-0.216‡	-0.085†	-0.098‡	-0.043	-0.106†
Workers commuting >2 hours	0.086	0.121	-0.046	0.007	0.081	0.002	-0.048
N	5,559	7,736	8,121	9,068	15,481	10,828	10,680
Adjusted R squared	0.429	0.461	0.366	0.446	0.418	0.442	0.356
Up to incomplete primary school							
Workers commuting 6'-30'	-0.035	0.018	0.029	0.008	0.001	0.048	-0.028
Workers commuting >30'-1 hour	-0.017	0.015	-0.033	0.050	0.033	0.084	0.008
Workers commuting >1-2 hours	-0.089	0.017	-0.129*	0.038	0.032	0.048	-0.054
Workers commuting >2 hours	0.062	0.016	-0.085	0.034	-0.136	-0.003	-0.038
N	1,754	2,889	2,777	3,918	4,158	4,804	3,017
Adjusted R squared	0.130	0.114	0.090	0.097	0.080	0.123	0.079
Complete primary school to high school graduates without college degree							
Workers commuting 6'-30'	-0.109†	-0.016	-0.056	-0.056	-0.016	-0.018	-0.029
Workers commuting >30'-1 hour	-0.084	-0.110*	-0.176‡	-0.078	-0.042	-0.105†	-0.014
Workers commuting >1-2 hours	-0.214‡	-0.163†	-0.211‡	-0.192‡	-0.096†	-0.113*	-0.121†
Workers commuting >2 hours	0.009	0.069	0.153	0.039	0.069	0.080	-0.022
N	3,015	3,979	4,187	4,091	8,071	4,749	6,550
Adjusted R squared	0.320	0.320	0.206	0.271	0.230	0.246	0.199
College degree							
Workers commuting 6'-30'	-0.171†	0.099	-0.058	0.041	-0.020	-0.051	0.063
Workers commuting >30'-1 hour	-0.367‡	0.057	-0.186	0.087	-0.116†	-0.115	-0.129
Workers commuting >1-2 hours	-0.452†	0.028	-0.293	-0.037	-0.286‡	-0.252	-0.192
Workers commuting >2 hours	0.104	0.579	-0.259	0.181	0.444†	-0.232	-0.128
N	790	868	1,157	1,059	3,252	1,275	1,113

Continued on next page

Table 5 – continued from previous page

Adjusted R squared	0.301	0.279	0.181	0.263	0.287	0.285	0.241

	Natal - RN	Grande Vitória - ES	Manaus - AM	Belém - PA	Goiânia - GO	Curitiba - PR	Forta-leza - CE
All individuals							
Workers commuting 6'-30'	-0.014	-0.027	0.029	-0.044	-0.018	0.011	-0.013
Workers commuting >30'-1 hour	-0.095‡	-0.075‡	-0.042	-0.066†	-0.089‡	-0.033*	-0.024
Workers commuting >1-2 hours	-0.109‡	-0.153‡	-0.121‡	-0.132‡	-0.193‡	-0.123‡	-0.112‡
Workers commuting >2 hours	0.078	-0.069*	-0.039	-0.015	-0.050	-0.102‡	-0.086*
N	12,056	24,887	11,912	15,523	16,951	32,523	27,034
Adjusted R squared	0.458	0.433	0.340	0.369	0.355	0.379	0.408
Up to incomplete primary school							
Workers commuting 6'-30'	-0.035	-0.095‡	0.045	-0.001	-0.005	0.006	0.031
Workers commuting >30'-1 hour	-0.023	-0.083†	0.048	-0.021	-0.028	0.014	0.091‡
Workers commuting >1-2 hours	-0.013	-0.125‡	-0.032	-0.106*	-0.122†	-0.018	0.038
Workers commuting >2 hours	-0.063	-0.054	-0.018	-0.077	0.017	-0.119†	-0.083
N	4,512	7,780	3,541	4,933	6,284	10,494	9,662
Adjusted R squared	0.122	0.096	0.089	0.074	0.062	0.088	0.078
Complete primary school to high school graduates without college degree							
Workers commuting 6'-30'	-0.002	0.004	-0.013	-0.061	-0.017	-0.032	-0.032
Workers commuting >30'-1 hour	-0.083†	-0.044	-0.119†	-0.049	-0.091†	-0.065‡	-0.052*
Workers commuting >1-2 hours	-0.162‡	-0.118‡	-0.172‡	-0.095†	-0.239‡	-0.176‡	-0.143‡
Workers commuting >2 hours	0.168	-0.041	-0.036	0.045	-0.105	-0.137†	-0.034
N	6,041	13,138	6,989	8,728	8,284	16,737	14,691
Adjusted R squared	0.262	0.219	0.186	0.184	0.210	0.180	0.214
College degree							
Workers commuting 6'-30'	-0.053	-0.059	0.143	-0.054	-0.060	0.128†	-0.137
Workers commuting >30'-1 hour	-0.359‡	-0.189‡	0.058	-0.223†	-0.227‡	-0.003	-0.289‡
Workers commuting >1-2 hours	-0.305†	-0.373‡	-0.217*	-0.360‡	-0.178	-0.249‡	-0.612‡
Workers commuting >2 hours	0.218	-0.187	0.004	-0.061	-0.135	0.323†	-0.323
N	1,503	3,969	1,382	1,862	2,383	5,292	2,681
Adjusted R squared	0.281	0.287	0.222	0.247	0.240	0.241	0.270

	Recife - PE	Salvador - BA	Porto Alegre - RS	Belo Horizonte - MG	Rio de Janeiro - RJ	São Paulo - SP
All individuals						
Workers commuting 6'-30'	-0.016	0.040	0.005	-0.013	-0.011	-0.018
Workers commuting >30'-1 hour	0.010	0.051†	-0.009	-0.048‡	-0.002	-0.014
Workers commuting >1-2 hours	-0.051†	0.042	-0.044†	-0.125‡	-0.031†	-0.067‡
Workers commuting >2 hours	-0.022	0.085†	-0.062	-0.134‡	-0.049†	-0.095‡
N	33,852	27,923	42,000	48,518	83,302	154,584
Adjusted R squared	0.406	0.409	0.424	0.419	0.397	0.367
Up to incomplete primary school						
Workers commuting 6'-30'	-0.050	0.020	-0.037	-0.068†	-0.082‡	-0.046†
Workers commuting >30'-1 hour	0.001	0.067	0.004	-0.035	-0.025	-0.012
Workers commuting >1-2 hours	-0.026	0.022	-0.028	-0.091‡	-0.030	-0.041†
Workers commuting >2 hours	-0.063	0.091	-0.110†	-0.092†	0.005	-0.047†
N	11,485	8,451	13,073	17,989	22,455	45,652
Adjusted R squared	0.094	0.084	0.092	0.078	0.072	0.077
Complete primary school to high school graduates without college degree						
Workers commuting 6'-30'	-0.010	0.018	0.004	-0.010	-0.017	-0.022
Workers commuting >30'-1 hour	0.023	0.031	-0.034	-0.058†	0.002	-0.030*
Workers commuting >1-2 hours	-0.040	0.038	-0.053†	-0.144‡	-0.032	-0.081‡
Workers commuting >2 hours	0.046	0.055	-0.032	-0.182‡	-0.075‡	-0.110‡
N	18,418	15,722	23,617	24,277	45,919	80,089
Adjusted R squared	0.201	0.211	0.222	0.212	0.181	0.175
College degree						
Workers commuting 6'-30'	0.072	0.150*	0.078	0.064	0.108†	0.022
Workers commuting >30'-1 hour	0.060	0.052	0.054	-0.020	0.025	0.026

Continued on next page

Table 5 – continued from previous page

Workers commuting >1-2 hours	-0.086	0.058	-0.067	-0.131†	-0.011	-0.053*
Workers commuting >2 hours	-0.162	0.173	0.034	0.007	-0.049	-0.162‡
N	3,949	3,750	5,310	6,252	14,928	28,843
Adjusted R squared	0.245	0.236	0.229	0.257	0.228	0.204

Source: Authors' calculations
Notes: Controls: age, age squared, colour or race, household head, with children up to 15 years old, married, sector of activity, occupation, existence of a formal contract. For the regressions with all individuals, the education attainment of the individual was included as an additional control. Reference category: workers commuting for up to 5 minutes. Significance levels: * p<0.10, †p<0.05, ‡p<0.01. Only male individuals aged 25 to 64 years old living within a distance of 30km from the centre are considered in the analysis. Sampling weights are taken into account with Stata command pweight. Complete tables can be requested from the authors.

Table 6: Logit model for the probability of being unemployed, regressions with all individuals and by education groups

	Macapá - AP	Aracaju - SE	Vale do Rio Cuiabá - MT	Maceió - AL	Florianó-polis - SC	João Pessoa - PB	Grande São Luís - MA
All individuals							
Inverse of distance	1.175	0.464	0.119	0.432	2.925†	1.410	1.713
Inverse of distance squared	0.944	6.556	31.848	1.390	0.455*	0.943	0.970
N	6,034	8,459	8,525	10,020	16,009	11,378	11,291
Pseudo R squared	0.055	0.064	0.038	0.053	0.051	0.067	0.061
Up to incomplete primary school							
Inverse of distance	0.322	0.192	0.003*	0.310	0.112	1.502	1.176
Inverse of distance squared	2.335	36.753	37,017.1*	1.407	4.366	0.912	0.992
N	1,917	3,223	2,952	4,448	4,320	5,083	3,186
Pseudo R squared	0.026	0.048	0.019	0.024	0.053	0.037	0.030
Complete primary school to incomplete tertiary school							
Inverse of distance	2.355	0.621	2.430	0.496	11.255	1.703	2.065*
Inverse of distance squared	0.584	2.704	0.058	1.339	0.138	0.889	0.960*
N	3,300	4,343	4,385	4,464	8,357	4,991	6,963
Pseudo R squared	0.064	0.047	0.038	0.056	0.062	0.068	0.061
Complete tertiary school							
Inverse of distance	0.531	98.561	0.620	1.195	2.783	0.172	1.181
Inverse of distance squared	1.648	0.061	37.412	0.749	0.484	1.852	0.992
N	817	893	1,188	1,108	3,332	1,304	1,142
Pseudo R squared	0.131	0.245	0.104	0.152	0.063	0.187	0.163

	Natal - RN	Grande Vitória - ES	Manaus - AM	Belém - PA	Goiânia - GO	Curitiba - PR	Forta-leza - CE
All individuals							
Inverse of distance	2.487	0.846	0.468	0.697	1.354	1.204	1.151
Inverse of distance squared	0.276	1.252	1.441	1.528	0.782	0.962	1.039
N	13,086	26,231	12,933	16,838	18,004	33,821	27,974
Pseudo R squared	0.053	0.035	0.035	0.038	0.030	0.026	0.046
Up to incomplete primary school							
Inverse of distance	2.789	0.124†	0.176*	0.257*	0.204	1.108	2.385
Inverse of distance squared	0.388	9.381*	2.620	3.555†	4.159	1.161	0.397
N	5,027	8,268	3,916	5,409	6,768	10,976	9,953
Pseudo R squared	0.038	0.026	0.016	0.021	0.030	0.022	0.026
Complete primary school to incomplete tertiary school							
Inverse of distance	1.332	1.989	0.734	1.010	4.190	3.190†	0.646
Inverse of distance squared	0.428	0.473	1.091	1.089	0.286*	0.706*	2.237
N	6,495	13,860	7,583	9,494	8,772	17,410	15,259
Pseudo R squared	0.050	0.031	0.033	0.038	0.032	0.026	0.051

Continued on next page

Table 6 – continued from previous page

Complete tertiary school							
Inverse of distance	1662.660*	3.270	0.542	1.987	0.448	0.117†	1.921
Inverse of distance squared	0.000	0.399	1.265	0.770	1.824	1.817†	0.278
N	1,564	4,103	1,434	1,935	2,464	5,435	2,762
Pseudo R squared	0.134	0.082	0.084	0.060	0.055	0.041	0.096

	Recife - PE	Salvador - BA	Porto Alegre - RS	Belo Horizonte - MG	Rio de Janeiro - RJ	São Paulo - SP
All individuals						
Inverse of distance	0.430†	0.411‡	2.409*	1.007	0.736	1.253
Inverse of distance squared	1.629	1.604†	0.694	0.848	1.176	0.654
N	37,419	28,533	37,203	49,194	84,152	164,255
Pseudo R squared	0.057	0.053	0.027	0.031	0.042	0.032
Up to incomplete primary school						
Inverse of distance	0.536	0.550	13.921*	0.569	0.275*	6.053‡
Inverse of distance squared	1.459	1.201	0.132‡	1.180	4.335	0.183‡
N	13,098	8,711	10,940	18,033	22,383	49,540
Pseudo R squared	0.027	0.023	0.027	0.022	0.018	0.016
Complete primary school to incomplete tertiary school						
Inverse of distance	0.342†	0.393†	2.030	0.865	0.877	0.644
Inverse of distance squared	1.732	1.722†	0.800	1.476	0.965	0.894
N	20,214	15,990	21,323	24,813	46,498	85,260
Pseudo R squared	0.053	0.050	0.024	0.031	0.041	0.029
Complete tertiary school						
Inverse of distance	1.012	0.143	0.803	1.430	54.645*	0.669
Inverse of distance squared	0.979	2.427	1.170	0.388	0.000	1.614
N	4,107	3,832	4,940	6,348	15,271	29,455
Pseudo R squared	0.090	0.080	0.028	0.032	0.055	0.028

Source: Authors' calculations
Notes: Controls: age, age squared, colour or race, household head, with children up to 15 years old, married. For the regressions with all individuals, the education attainment of the individual was included as an additional control. Coefficients are presented as odds-ratios. Significance levels: * p<0.10, †p<0.05, ‡p<0.01. Only male individuals aged 25 to 64 years old living within a distance of 30km from the centre are considered in the analysis. Sampling weights are taken into account with Stata command pweight. Complete tables can be requested from the authors.

A few aspects can be highlighted in relation to these results. On the one hand, unemployment levels may vary throughout the city in an irregular way, with no specific pattern in either monocentric or multicentric cities. In a sense, this conclusion in the Brazilian case matches part of the literature, which finds no regular pattern for the spatial distribution of the unemployment rate.

However, the conclusion goes against recent theoretical predictions that distance to jobs can affect the probability that individuals belonging to low-skilled minorities find a position. If these theoretical predictions are valid, it might be that there are methodological issues driving this unexpected result. First, distance is not measured in relation to an individual, but relates only to her neighbourhood. In addition, we do not take into account the location of job offers and existing jobs. Our database locates individuals by their place of residence. Therefore, there may be difficulties in correctly identifying the centres in the city and in calculating the relative location of each potential worker. Moreover, when distance is measured as the commuting time for workers in the neighbourhood, this may not be the same as the commuting time a potential worker would spend if he or she were in work.

Table 7: Logit model for the probability of being unemployed, regressions with all individuals and by education groups

	Macapá - AP	Aracaju - SE	Vale do Rio Cuiabá - MT	Maceió - AL	Florianó-polis - SC	João Pessoa - PB	Grande São Luís - MA
All individuals							
% workers commuting 6'-30'	0.048†	0.664	-0.091	0.293	0.327	0.212	0.507
% workers commuting >30'-1 hour	0.105	0.288	2,219	0.175	0.049†	0.475	0.168†
% workers commuting >1 hour	0.516	0.929	2.128*	1,205	0.428	1,267	0.723
N	6,034	8,459	8,525	10,020	16,009	11,782	11,566
Adjusted R squared	0.056	0.063	0.040	0.053	0.054	0.064	0.061
Up to incomplete primary school							
% workers commuting 6'-30'	0.011*	3,231	-0.161	1,304	0.004†	0.151	1,532
% workers commuting >30'-1 hour	0.005*	0.995	5,118	1,424	0.021*	0.159	0.546
% workers commuting >1 hour	1,398	1,639	2,514	1,715	0.006‡	0.898	1,071
N	1,917	3,223	2,952	4,448	4,320	5,389	3,291
Adjusted R squared	0.026	0.046	0.019	0.025	0.059	0.035	0.030
Complete primary school to high school graduates without college degree							
% workers commuting 6'-30'	0.101	0.299	0.040	0.047†	2,798	0.799	0.291
% workers commuting >30'-1 hour	0.336	0.160	3,518	0.018†	0.078	3,474	0.081†
% workers commuting >1 hour	0.349	0.884	2,588	0.631	1,368	3,524	0.570†
N	3,300	4,343	4,385	4,464	8,357	5,079	7,122
Adjusted R squared	0.064	0.047	0.039	0.058	0.067	0.067	0.063
College degree							
% workers commuting 6'-30'	0	0.012	-0.301	0.337	0.210	0.000	2,668
% workers commuting >30'-1 hour	0.056	0.005	0.031	0.010	0.061	0.002	0.574
% workers commuting >1 hour	0.158	0.003†	0.240	26.935*	4,475	0.002	0.999
N	817	893	1,188	1,108	3,332	1,314	1,153
Adjusted R squared	0.139	0.260	0.102	0.177	0.064	0.162	0.164

	Natal - RN	Grande Vitória - ES	Manaus - AM	Belém - PA	Goiânia - GO	Curitiba - PR	Forta-leza - CE
All individuals							
% workers commuting 6'-30'	0.196*	3,722	0.322	0.869	0.274	1,508	8.937*
% workers commuting >30'-1 hour	0.111†	5,284	5,128	0.357	1,030	1,296	5,723
% workers commuting >1 hour	0.790	1.715*	0.847	2,357	0.756	1,576	4,325
N	13,086	26,231	12,933	16,838	17,626	33,594	28,821
Adjusted R squared	0.054	0.035	0.037	0.039	0.032	0.026	0.046
Up to incomplete primary school							
% workers commuting 6'-30'	0.126	1,701	0.493	36,757	0.210	0.816	2,237
% workers commuting >30'-1 hour	0.257	2,175	5,536	5,896	0.698	0.868	0.353
% workers commuting >1 hour	0.705	1,634	1,028	132.388*	1,024	1,929	3,518
N	5,027	8,268	3,916	5,409	6,579	10,862	10,416
Adjusted R squared	0.038	0.025	0.016	0.022	0.033	0.022	0.028
Complete primary school to high school graduates without college degree							
% workers commuting 6'-30'	0.155	13.582†	0.166	0.102	0.496	2,746	113.524‡
% workers commuting >30'-1 hour	0.038†	14.542†	4,042	0.074	1,371	2,788	110.358‡
% workers commuting >1 hour	0.711	2.230*	0.657	0.223	0.741	1,232	19.355†
N	6,495	13,860	7,583	9,494	8,604	17,310	15,623
Adjusted R squared	0.051	0.032	0.036	0.039	0.032	0.025	0.051
College degree							
% workers commuting 6'-30'	0.343	0.047	13,079	0.266	0.234	0.305	0.042
% workers commuting >30'-1 hour	0.493	1,331	901,885	0.139	10,454	0.064	4,959
% workers commuting >1 hour	0.524	0.180	1,656	0.878	0.071†	3,154	0.004
N	1,564	4,103	1,434	1,935	2,443	5,422	2,782
Adjusted R squared	0.128	0.086	0.086	0.059	0.065	0.038	0.102

Continued on next page

Table 7 – continued from previous page

	Recife - PE	Salvador - BA	Porto Alegre - RS	Belo Horizonte - MG	Rio de Janeiro - RJ	São Paulo - SP
All individuals						
% workers commuting 6'-30'	0.545	0.503	0.360	0.487	1,264	0.679
% workers commuting >30'-1 hour	0.982	0.264	0.433	0.901	0.909	0.771
% workers commuting >1 hour	1,370	1,047	0.376†	0.733	1,158	1,020
N	37,669	30,873	43,722	50,753	88,531	164,684
Adjusted R squared	0.057	0.051	0.024	0.031	0.042	0.032
Up to incomplete primary school						
% workers commuting 6'-30'	0.273	1,333	1,986	0.278	0.935	0.628
% workers commuting >30'-1 hour	0.609	0.358	1,593	1,120	0.696	0.779
% workers commuting >1 hour	0.699	2,183	1,376	0.738	1,321	0.981
N	13,207	9,678	13,696	18,924	24,195	49,331
Adjusted R squared	0.027	0.023	0.022	0.024	0.019	0.016
Complete primary school to high school graduates without college degree						
% workers commuting 6'-30'	0.670	0.404	0.067†	0.358	1,048	0.902
% workers commuting >30'-1 hour	1,243	0.249	0.111†	0.396	0.948	0.873
% workers commuting >1 hour	1,854	0.837	0.121‡	0.526	1,049	1,094
N	20,350	17,292	24,572	25,418	48,893	85,450
Adjusted R squared	0.053	0.048	0.023	0.031	0.042	0.029
College degree						
% workers commuting 6'-30'	6,864	0.015	17,634	12,783	0.689	0.337
% workers commuting >30'-1 hour	1,456	1,056	18,005	55,136	0.345	0.512
% workers commuting >1 hour	7,570	0.148	2,771	2,117	0.523	0.915
N	4,112	3,903	5,454	6,411	15,443	29,903
Adjusted R squared	0.091	0.080	0.024	0.030	0.058	0.029

Source: Authors' calculations
Notes: Controls: age, age squared, colour or race, household head, with children up to 15 years old, married. For the regressions with all individuals, the education attainment of the individual was included as an additional control. Coefficients are presented as odds-ratios. Significance levels: * p<0.10, †p<0.05, ‡p<0.01. Only male individuals aged 25 to 64 years old living within a distance of 30km from the centre are considered in the analysis. Sampling weights are taken into account with Stata command pweight. Complete tables can be requested from the authors.

As robustness checks, some additional results are provided in Table A.4, in the appendix[5]. We run the models for all individuals without dividing the database between metropolitan areas. Then, we also include in these models the control for the metropolitan area of residence. As it can be seen, the regression for the logarithm of the hourly wage against distance measures indicate that the farther away from the city centre the lower the wage received on average. A higher education attainment is associated to higher wages, which are also present in some of the largest metropolitan areas. However, when Model 2 is considered, the partial correlation of commuting distance and wages does not have the expected sign (longer commuting should be associated with lower wages). This is due to the fact that longer time periods are more common in larger urban areas, which are associated to more populated metropolitan areas. This is an indication that there are iterative effects of commuting distance and city size which should be controlled for (what is done in the estimations presented in Table 5).

For the models related to the probability of unemployment, the results indicate that a lower chance of unemployment is associated with longer commuting times in the weighting area (once again, an unexpected result). This is most likely caused by the fact that iterations between metropolitan areas and commuting times are not taken into account. In addition, the probability of unemployment is lower for more educated individuals. These considerations make our previous estimations preferable in relation to this additional exercise.

[5]We thank the contribution of an anonymous referee who suggested that we estimated these alternative models to enrich our analysis.

5 Final remarks

There is significant spatial mismatch in the labour market in Brazilian metropolitan areas. The influence of spatial location and distance to jobs on labour market outcomes is stronger for larger urban areas, and wages are more strongly related to distance to jobs and to distance to the centre than unemployment rates are. In addition, the difference in the commuting time for poor and rich workers is larger in labour markets with 500,000 workers or more.

The literature on spatial mismatch suggests that this phenomenon is predominantly urban and that it is more relevant for low-skilled minorities in larger urban areas for whom congestion costs are relatively more important. In addition, these minorities may face more limitations in their social interactions, with a significant impact on their ability to find a better match in the job market.

In this paper, we have attempted to investigate whether this negative relationship between spatial mismatch and labour market outcomes is valid in Brazil after controlling for individual characteristics. Our conclusions indicate that there is no clear relation between two different measures of accessibility to jobs and the probability of being unemployed. However, for wages there is a clear correlation, which is stronger in larger metropolitan areas.

These results indicate that in the Brazilian case, the spatial mismatch is more relevant to determine individual wages (in accordance to the relationship mentioned by Gobillon et al. 2007). On the other hand, the probability of unemployment may not be affected as much by it. This can be a result of the empirical strategy adopted here, in which commuting time spent by workers is used to calculate the potential commuting time an unemployed person would have spent in case she was employed. It may also be an indication that the spatial mismatch has a stronger effect than alternative measures of unbalance in the labour market, such as unemployment duration, as it was found in the literature (see for instance Rogers 1997). Finally, the adequate estimation strategy should allow for iterative effects between accessibility measures and metropolitan areas. This means that each metropolitan area has a particular dynamic in the labour market.

In any case, city size and skill level seem to be relevant aspects for the chances an individual has to perform well in the labour market. Intra-urban policies should aim to reduce inequalities in terms of accessibility. Since education attainment is strongly related to income, poorer neighbourhoods, which are also less served by public policies – and in the peripheries are usually far away from jobs – should be the main focus of transportation policies in the short run and education programs for middle and long run results.

This is intended to be an exploratory work. In this sense, we have explored correlations between labour market outcomes and measures of accessibility to jobs for Brazilian metropolitan areas. Our results depend on strong identification hypotheses to avoid bias related to simultaneous location decisions of workers and firms within the city (Ihlanfeldt 2006). If these conditions do not hold, our results may not represent a causal relationship, but will be meaningful in the sense of providing a better understanding of the conditional distribution of wages and the unemployment rate in the biggest metropolitan areas of Brazil.

The broader analysis of urban labour markets in Brazil provides an indication that there are relevant differences in the way workers and firms interact in space, and urban scale seems to be important to this relationship. Future work should investigate these issues more thoroughly. In this sense, different proximity dimensions could be included in the analysis, in order to investigate the factors that generate the spatial mismatch. However, this approach would require a more comprehensive database of the characteristics of Brazilian labour markets and the local interaction between individuals, which are not available yet.

Acknowledgement

Eduardo Amaral Haddad acknowledges financial support from CNPq.

References

Andersson F, Haltiwanger JC, Kutzbach MJ, Pollakowsky HO, Weinberg DH (2014) Job displacement and the duration of joblessness: The role of spatial mismatch. NBER working paper series, WP 20066. CrossRef.

Åslund O, Östh J, Zenou Y (2010) How important is access to jobs? Old question – improved answer. *Journal of Economic Geography* 10: 389–422. CrossRef.

Brueckner JK, Thisse JF, Zenou Y (2002) Local labor markets, job matching, and urban location. *International Economic Review* 43[1]: 155–171. CrossRef.

Bruzelius N (1979) *The Value of Travel Time*. Croom Helm, London

Capello R (2007) *Regional Economics*. Routledge Advanced Texts in Economics and Finance, Abingdon

Chauvin JP, Glaeser E, Ma Y, Tobio K (2016) What is the difference about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. *Journal of Urban Economics* 98: 17–49. CrossRef.

Da Mata D, Deichmann U, Henderson JV, Lall SV, Wang HG (2007) Determinants of city growth in Brazil. *Journal of Urban Economics* 62: 252–272. CrossRef.

Di Paolo A, Matas A, Raymond JL (2016) Job accessibility and job-education mismatch in the metropolitan area of Barcelona. *Papers in Regional Science* 96[S1]: S91–S112. CrossRef.

Fujita M, Thisse JF (2012) *Economics of Agglomeration: Cities, industrial location and regional growth* (2nd ed.). Cambridge University Press, Cambridge

Gibb K, Osland L, Pryce G (2014) Describing inequalities in access to employment and the associated geography of wellbeing. *Urban Studies* 51[3]: 596–613. CrossRef.

Gobillon L, Selod H (2013) Spatial mismatch, poverty, and vulnerable populations. In: Fischer MM, Nijkamp P (eds), *Handbook of Regional Science*. Springer, 93-107. CrossRef.

Gobillon L, Selod H, Zenou Y (2007) The mechanisms of spatial mismatch. *Urban Studies* 44[12]: 2401–2427. CrossRef.

Haddad EA, Barufi AMB (2017) From rivers to roads: Spatial mismatch and inequality of opportunity in urban labor markets of a megacity. *Habitat International* 68: 3–14. CrossRef.

Holzer HJ (1991) The spatial mismatch hypothesis: What has the evidence shown? *Urban Studies* 28[1]: 105–122. CrossRef.

Houston D (2005a) Employability, skills mismatch and spatial mismatch in metropolitan labour markets. *Urban Studies* 42[2]: 221–243. CrossRef.

Houston D (2005b) Methods to test the spatial mismatch hypothesis. *Economic Geography* 81[4]: 407–434. CrossRef.

Hu L (2014) Changing job access of the poor: Effects of spatial and socioeconomic transformations in Chicago. *Urban Studies* 51[4]: 675–692

Hu L, Giuliano G (2014) Poverty concentration, job access, and employment outcomes. *Journal of Urban Affairs* 39[1]: 1–16. CrossRef.

Ihlanfeldt KR (2006) A primer on spatial mismatch within urban labor markets. In: Arnott RJ, McMillen DP (eds), *A Companion to Urban Economics*. 404-417, Oxford. CrossRef.

Johnson RC (2006) Landing a job in urban space: The extent and effects of spatial mismatch. *Regional Science and Urban Economics* 36: 331–372. CrossRef.

Kain JF (1968) Housing segregation, negro employment, and metropolitan decentralization. *Quarterly Journal of Economics* 82: 32–59. CrossRef.

Kain JF (1992) The spatial mismatch hypothesis: Three decades later. *Housing Policy Debate* 3[2]: 371–459. CrossRef.

Kain JF (2004) A pioneer's perspective on the spatial mismatch literature. *Urban Studies* 41[1]: 7–32. CrossRef.

Krugman P (1995) *Development, Geography, and Economic Theory*. The MIT Press, Cambridge, Mass

Lucas Jr RE, Rossi-Hansberg E (2002) On the internal structure of cities. *Econometrica* 70[4]: 1445–1476. CrossRef.

McCann P (2013) *Modern Urban and Regional Economics* (2nd ed.). Oxford University Press, Oxford

Morrison PS (2005) Unemployment and urban labour markets. *Urban Studes* 42[1]: 2261–2288. CrossRef.

Ottaviano G (2004) Chapter 58 – Agglomeration and economic geography. In: Henderson JV, Thisse JF (eds), *Handbook of Regional and Urban Economics, Volume 4*. Elsevier, Amsterdam, 2563–2608

Partridge M, Rickman DS, Ali K, Rose-Olfert M (2009) Agglomeration spillovers and wage and housing cost gradients across the urban hierarchy. *Journal of International Economics* 78: 126–140. CrossRef.

Roback J (1982) Wages, rents and the quality of life. *Journal of Political Economy* 90[6]: 1257–1278. CrossRef.

Rogers CL (1997) Job search and unemployment duration: Implications for the spatial mismatch hypothesis. *Journal of Urban Economics* 42: 109–132. CrossRef.

Topa G, Zenou Y (2015) Neighborhood and network effects. In: Duranton G, Henderson JV, Strange WC (eds), *Handbook of Regional and Urban Economics, Volume 5A*. Elsevier, Amsterdam, 561–624. CrossRef.

Tyndall J (2015) Waiting for the R train: Public transportation and employment. *Urban Studies* 54[2]: 520–537. CrossRef.

Vieira R, Haddad EA (2015) An accessibility index for the metropolitan area of São Paulo. In: Kourtit K, Nijkamp P, Stough RR (eds), *The Rise of the City: Spatial Dynamics in the Urban Century*. Edward Elgar Publishing, Cheltenham, UK, 242–258. CrossRef.

Zenou Y (1999) Unemployment in cities. In: Huriot JM, Thisse JF (eds), *Economics of Cities*. Cambridge University Press, Cambridge, 343–389

Zenou Y (2000) Urban unemployment, agglomeration and transportation policies. *Journal of Public Economics* 77: 97–133. CrossRef.

Zenou Y (2006) Efficiency wages and unemployment in cities: The case of high-relocation costs. *Regional Science and Urban Economics* 36: 49–71. CrossRef.

Zenou Y (2009) *Urban Labor Economics* (1st ed.). Cambridge University Press, New York. CrossRef.

Zenou Y (2013) Spatial versus social mismatch. *Journal of Urban Economics* 74: 113–132. CrossRef.

A Appendix

Table A.1: Average hourly wage (in Brazilian real) in each weighting area by the distance to the main business centre, 2010.

	Distance from centre (in kilometer)							
	<2.5	2.5 to <5	5 to <10	10 to <20	20 to <30	30 to <40	40 to <50	50 or more
Macapá - AP	12.96	10.98	10.00	7.61		8.47		7.32
Aracaju - SE	11.80	14.75	9.08	9.07	5.29			
Vale do Rio Cuiabá - MT	13.95	24.74	11.97	8.83	8.03			6.57
Maceió - AL	19.11	6.60	10.14	6.53	5.28	4.90		
Florianópolis - SC	24.60	20.15	15.59	11.07	9.30	8.31	7.71	7.64
João Pessoa - PB	7.58	10.99	12.01	7.04	3.72	3.91	4.65	4.48
Grande São Luís - MA	8.79	14.08	13.28	9.52	5.33	4.58		
Natal - RN	5.60	6.14	11.04	13.95	4.62	5.64	5.24	5.24
Grande Vitória - ES	13.59	15.30	12.84	10.00	8.10	7.79	10.41	10.41
Manaus - AM	12.44	14.44	15.31	9.42			4.58	6.68
Belém - PA	20.14	13.85	11.42	8.44	6.19	5.32	5.73	5.73
Goiânia - GO	28.69	18.71	13.19	9.21	6.69	7.12	7.21	9.00
Curitiba - PR	28.58	31.99	14.51	9.45	9.10	6.62	7.38	6.68
Fortaleza - CE	6.13	7.67	12.46	10.37	6.73	4.44	3.84	4.19
Salvador - BA	12.63	23.79	9.69	8.68	10.41	7.32	8.89	7.62
Recife - PE	21.29	12.68	11.67	8.25	6.46	4.66	6.67	6.38
Porto Alegre - RS	26.99	31.46	17.78	10.09	9.15	8.65	9.05	8.41
Belo Horizonte - MG	34.98	19.37	14.90	9.85	7.50	8.90	6.88	7.79
Rio de Janeiro - RJ	7.74	15.58	17.51	14.84	11.77	8.91	8.08	8.25
São Paulo - SP	23.47	21.54	21.51	16.88	10.69	11.14	10.09	9.96

Source: IBGE

Table A.2: Average individual hourly wage (in Brazilian real) by commuting time from home to work, 2010.

	Up to 5 min.	6 min to $\frac{1}{2}$ hour	$>\frac{1}{2}$ to 1 hour	>1 to 2 hours	>2 hours
Macapá - AP	10.96	11.04	8.32	7.39	12.36
Aracaju - SE	11.72	12.16	9.27	7.85	16.58
Vale do Rio Cuiabá - MT	18.88	14.47	11.81	7.75	15.31
Maceió - AL	7.96	10.01	9.21	6.87	11.14
Florianópolis - SC	13.52	14.19	13.48	11.10	15.77
João Pessoa - PB	9.56	10.97	7.99	6.56	6.93
Grande São Luís - MA	11.16	12.48	10.49	7.76	10.94
Natal - RN	10.81	11.35	7.66	6.61	11.45
Grande Vitória - ES	14.40	13.23	11.05	8.31	10.10
Manaus - AM	10.01	13.60	10.38	7.82	8.81
Belém - PA	12.18	11.49	10.70	7.99	12.06
Goiânia - GO	17.91	13.44	9.99	7.42	14.08
Curitiba - PR	14.54	15.32	12.55	8.54	9.45
Fortaleza - CE	10.04	10.45	8.80	6.33	8.43
Salvador - BA	9.45	10.85	11.23	11.08	13.33
Recife - PE	10.58	10.17	10.50	8.36	8.29
Porto Alegre - RS	12.27	13.50	11.55	9.82	9.97
Belo Horizonte - MG	13.18	13.02	11.55	9.25	9.05
Rio de Janeiro - RJ	14.22	12.87	13.29	12.77	10.46
São Paulo - SP	17.53	16.79	15.83	13.15	11.95

Source: IBGE

Table A.3: Average unemployment rate in each weighting area by the distance to the main business centre, 2010

	Distance from centre (in kilometer)							
	<2.5	2.5 to <5	5 to <10	10 to <20	20 to <30	30 to <40	40 to <50	50 or more
Macapá - AP	8.4%	6.7%	6.6%	10.3%		12.4%		4.7%
Aracaju - SE	10.0%	6.6%	6.9%	9.5%	7.3%			
Vale do Rio Cuiabá - MT	4.1%	3.2%	4.4%	4.3%	7.7%			8.9%
Maceió - AL	5.4%	8.6%	7.5%	8.3%	13.0%	11.8%		
Florianópolis - SC	2.5%	3.8%	2.7%	1.8%	3.4%	3.9%	1.6%	1.9%
João Pessoa - PB	9.1%	5.3%	5.5%	6.8%	8.0%	8.0%	13.9%	8.8%
Grande São Luís - MA	7.7%	10.1%	7.2%	7.7%	6.0%	4.1%		
Natal - RN	7.3%	8.0%	7.4%	5.9%	8.7%	5.4%	8.6%	
Grande Vitória - ES	5.9%	4.5%	4.4%	4.9%	5.7%	4.9%	7.3%	
Manaus - AM	6.0%	8.2%	6.4%	7.7%			4.2%	6.9%
Belém - PA	6.8%	7.4%	5.3%	7.3%	7.6%	5.8%	8.1%	
Goiânia - GO	3.5%	3.2%	3.0%	3.6%	3.5%	4.4%	3.8%	5.8%
Curitiba - PR	3.4%	2.8%	3.0%	2.9%	3.4%	1.5%	2.2%	3.3%
Fortaleza - CE	5.2%	6.8%	5.4%	5.2%	6.7%	6.9%	5.1%	6.3%
Salvador - BA	8.6%	5.7%	8.2%	9.5%	9.3%	12.2%	12.9%	14.6%
Recife - PE	6.7%	7.9%	9.2%	9.7%	11.2%	11.6%	9.6%	11.9%
Porto Alegre - RS	3.7%	4.4%	4.0%	3.7%	3.9%	3.6%	3.1%	2.7%
Belo Horizonte - MG	2.7%	3.8%	4.7%	4.3%	4.1%	4.2%	4.7%	3.7%
Rio de Janeiro - RJ	6.0%	5.1%	5.1%	5.3%	6.0%	6.3%	6.7%	7.8%
São Paulo - SP	5.5%	5.2%	5.1%	5.9%	6.1%	5.5%	5.0%	4.1%

Source: IBGE

Table A.4: Regressions for the whole database

	Model 1 ln(hourly wage) OLS	Model 2 ln(hourly wage) OLS	Model 3 unemp. (P = 1) Logit	Model 4 unemp. (P = 1) Logit
Inverse of distance	1.177‡		1.024	
Inverse of distance squared	0.990‡		0.999	
Commuting time of workers in the weighting area				
% workers commuting 6' to 30'				0.547†
% workers commuting more than 30' to 1 hour				0.543†
% workers commuting more than 1 hour				0.876
Individual commuting time (Reference: up to 5')				
6' to 30'		0.994		
More than 30' to 1 hour		0.979‡		
More than 1 hour to 2 hours		0.930‡		
More than 2 hours		0.932‡		
Metropolitan area (Reference: Belém - PA, 402,170 men 25-64)				
Macapá - AP (85,494 men 25-64)	0.996	1.007	1.078	1.093
Aracaju - SE (159,838 men 25-64)	0.935‡	0.932‡	1.168‡	1.198‡
Vale do Rio Cuiabá - MT (160,638 men 25-64)	1.154‡	1.146‡	0.647‡	0.666‡
Maceió - AL (216,904 men 25-64)	0.842‡	0.842‡	1.323‡	1.334‡
Florianópolis - SC (217,208 men 25-64)	1.186‡	1.175‡	0.439‡	0.447‡
João Pessoa - PB (230,930 men 25-64)	0.811‡	0.826‡	0.995	1.026
Grande São Luís - MA (244,017 men 25-64)	0.997	0.987	1.134†	1.138†
Natal - RN (258,207 men 25-64)	0.878‡	0.872‡	1.150‡	1.171‡
Grande Vitória - ES (353,561 men 25-64)	1.107‡	1.102‡	0.804‡	0.800‡
Manaus - AM (378,496 men 25-64)	1.106‡	1.105‡	1.078	1.073
Goiânia - GO (415,541 men 25-64)	1.150‡	1.139‡	0.547‡	0.550‡
Curitiba - PR (623,103 men 25-64)	1.167‡	1.157‡	0.523‡	0.525‡
Fortaleza - CE (666,504 men 25-64)	0.875‡	0.869‡	0.874‡	0.878‡
Salvador - BA (723,297 men 25-64)	0.989	0.987	1.309‡	1.279‡
Recife - PE (745,952 men 25-64)	0.862‡	0.856‡	1.522‡	1.512‡
Porto Alegre - RS (807,268 men 25-64)	1.081‡	1.068‡	0.627‡	0.629‡
Belo Horizonte - MG (1,115,715 men 25-64)	1.105‡	1.097‡	0.679‡	0.665‡
Rio de Janeiro - RJ (2,402,075 men 25-64)	1.131‡	1.122‡	0.916†	0.865‡
São Paulo - SP (3,953,270 men 25-64)	1.236‡	1.225‡	1.006	0.947
Education attainment (Reference: up to incomplete primary school)				
Complete primary school to incomplete college	1.304‡	1.305‡	0.782‡	0.788‡
Complete college	2.576‡	2.588‡	0.426‡	0.440‡
N	583,184	583,184	621,359	621,359
Pseudo R squared			0.048	0.048
Adjusted R squared	0.406	0.405		

Source: Authors' calculations
Notes: Controls for Models 1 and 2: age, age squared, colour or race, household head, with children up to 15 years old, married, sector of activity, occupation, existence of a formal contract. Controls for Models 3 and 4: age, age squared, colour or race, household head, with children up to 15 years old, married. Coefficients are presented as odds-ratios. Significance levels: * p <0.10, †p <0.05, ‡p <0.01. Only male individuals aged 25 to 64 years old living within a distance of 30km from the centre are considered in the analysis. Sampling weights are taken into account with Stata command pweight. Complete tables are available under request to the authors.

Resources

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 3, 2017, R1–R6
DOI: 10.18335/region.v4i3.198

journal homepage: region.ersa.org

The Chilean Internal Migration (CHIM) database: A Temporally Consistent Spatial Data Framework for the Analysis of Human Mobility

Francisco Rowe[1]

[1] University of Liverpool, Liverpool, UK (email: F.Rowe-Gonzalez@liverpool.ac.uk)

Received: 20 July 2017/Accepted: 22 July 2017

Abstract. Changes in zonal boundaries and coding schemes severely compromise temporal comparison of data. In Chile, the Population and Housing census is the only comprehensive source of internal migration data, but municipal boundaries and occupation and industry sector coding schemes have undergone various changes which hamper the temporal comparability of census data. This paper presents the Chilean Internal Migration database which provides a temporally consistent framework for the analysis of internal migration over a period covering twenty-five years from 1977 to 2002. Specifically, it offers a hierarchical system of 304 municipalities, 51 provinces, and 13 regions, with 10 occupational groups and 11 industrial sectors which are temporally consistent over the 1977-82, 1987-92 and 1997-2002 census intervals. The database can be downloaded from: https://s3.amazonaws.com/geoda/data/CHIM.zip.

1 Background

Spatio-temporal analysis presumes temporally consistent geographical and data coding frameworks. However, data sources are commonly affected by changes in geographical boundaries and by changes in the way information is registered. In Chile, the Population and Housing census represents the most comprehensive source of internal migration and socio-economic data. The Chilean censuses conducted in 1982, 1992 and 2002 were all affected by changes in municipal boundaries and various schemes have been used to classify information on education, industry and occupation.

1.1 Changes to census data

Internal migration is inherently a spatial process so changes in municipal boundaries present a challenge for its temporal analysis. In Chile, municipalities represent the smallest geographical unit to which the place of usual residence is recorded and internal migration data can be derived. Any alteration to boundaries directly compromises the comparability of migration statistics over time. Various changes affected municipal boundaries over the three censuses (Rowe 2013): between the 1982 census and 1992 census, 82 of out 335 municipalities underwent boundary shifts, affecting 2.2% of the total population and 247 thousand people; while between 1992 and 2002 censuses, only ten municipalities were affected – but these changes involved a much larger number of people (321 thousand), increasing the total number of municipalities from 335 in 1992 to 342 in 2002 due to the division of existing areas and creation of new municipalities.

Additionally, census records also underwent changes in coding classification schemes. Different classification systems have been employed to organize information on occupations, industry sectors, and education categories for each census. Occupations were recorded using the Americas Classification of Occupations (COTA-70-developed by the Statistical Institute of American States (IASI)) at the 1982 census but this system was replaced for the International Standard Classification of Occupations (ISCO-88 -developed by the International Labour Organization (ILO)) at the 1992 and 2002 censuses. To classify data on industry sectors, the International Standard Industrial Classification of All Economic Activities (ISIC) Revision 2 was used at the 1982 census – which was then replaced for the ISIC Revision 3 for the 1992 and 2002 censuses. To record information on education, different national classification systems designed by the Chilean National Statistical Institute (INE) have been used for each of the three censuses.

1.2 Harmonization of census data

All these changes hinder spatio-temporal analysis of census data. To enable this, Rowe, Bell (2013) built a database, labelled Chilean Internal Migration (CHIM) database, based on census micro-data by harmonizing municipal boundaries and using consistent coding schemes. Census micro-data were obtained from the INE. To harmonize municipal boundary changes, Rowe, Bell (2013) used two procedures. For boundary shifts that occurred between 1982 and 1992, they implemented a 'pseudo construct designer zones' approach. It involves the construction of purpose-built zones from smaller building-bricks to harmonize the zonal system from different census periods based on a common set of boundaries. For boundary changes that occurred between 1992 and 2002, Rowe, Bell (2013) used a 'freeze history' approach. This approach consists of freezing the zonal system at a certain point in time and systematically tracks subsequent alterations in geographical boundaries. In doing so, later observations can be adjusted back to the original geography. For details on these procedures, see Rowe, Bell (2013).

To harmonize education, occupation and industry sector data, the United Nations' international standard classification systems were used. Education data were recoded into five categories based on the International Standard Classification of Education, ISCED-1997 developed by The United Nations Educational Scientific and Cultural Organization (UNESCO 1997). To classify industry and occupation data, the ISIC Rev. 3 and ISCO-88 (respectively), which were used for the 1992 and 2002 censuses, were employed. Data from the 1982 census were mapped onto these schemes and industry and occupation information for all three censuses was then organized into 11 industry sectors and 10 occupational groups.

1.3 CHIM database

The outcome from the above procedures is a temporally consistent geography comprising 304 municipalities, 51 provinces, and 13 regions (Figure 1). The resulting municipal areas were defined as Temporal Municipalities (TMs) and the resulting hierarchical zonal system was classified using the INE's current geographical classification system. Regions were numbered using the conventional order from 01 to 13 from Tarapacá to the Metropolitan Region. Provinces were numbered incrementally within each region using a three-digit identifier. So, for example, provinces within Tarapacá were numbered 011, 012 and 013, and within Los Lagos 101, 102, 103, 104 and 105. TMs were also numbered incrementally within provinces using a five-digit identifier. So, for example, TMs within the province of Iquique were numbered from 01101, 01102, and 01103 up to 01106, and within Llanquihue from 10101 up to 10108. The original municipality names at the 1992 census were preserved, and in cases where municipalities were amalgamated, the name of the municipality with the largest population in 1992 was adopted.

The CHIM database covers the full population at each census year, 1982, 1992 and 2002, and is a composite structure of multiple data sets: a set of micro-data, and a set of four aggregate data files[1]. Figure 2 shows the aggregate files which consist of (1)

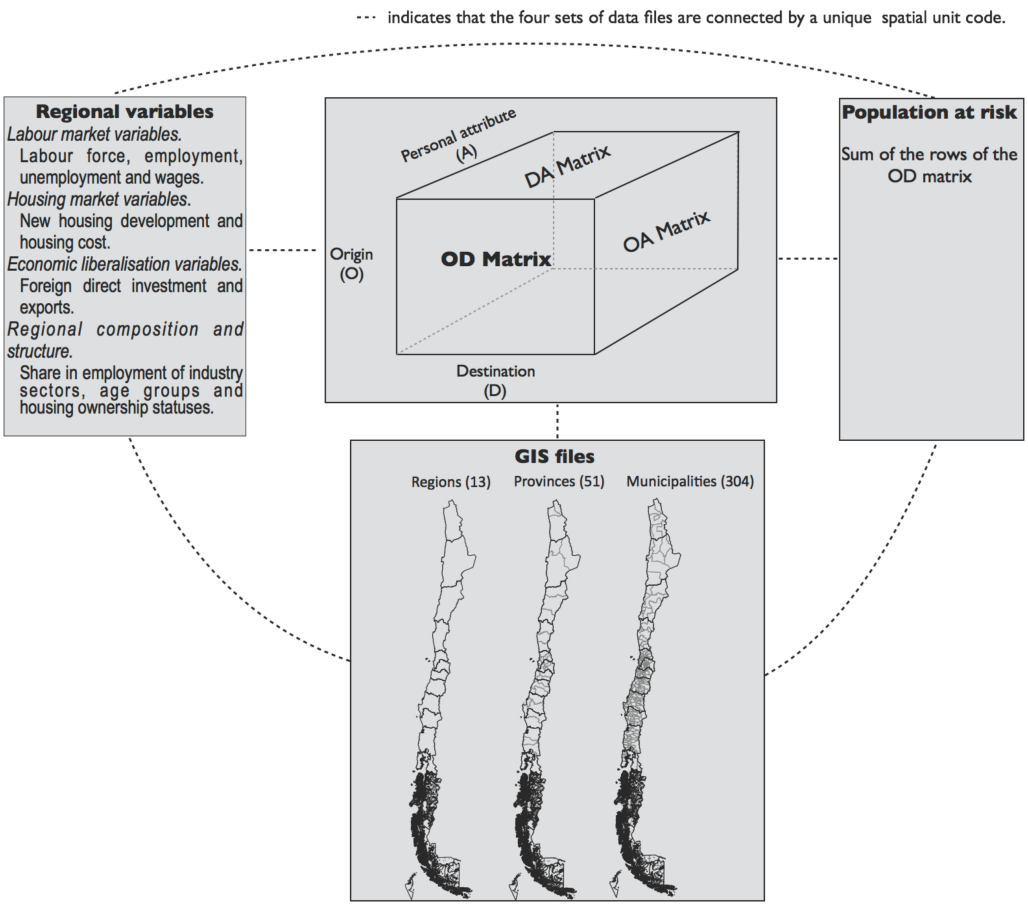
[1]Data from the 2012 Chilean census are not included in the CHIM database. This census has been questioned due to problems of imputation of unobserved housing units, with no identification of imputed

Notes: TAR: Tarapacá, ANT: Antofagasta, ATA: Atacama, COQ: Coquimbo, VAL: Valparaíso, MR: Metropolitan Region, OHI: O'higgins, MAU: Del Maule, BIO: Bío-bío, ARA: Araucanía, LAG: Los Lagos, AIS: Aísen, MAG: Magallanes. A full list of province and municipality names and codes is available: https://s3.amazonaws.com/geoda/data/CHIM.zip

Figure 1: Temporally consistent geography for Chile

origin-destination migration matrices, (2) populations at risk, (3) digital boundaries, and (4) regional contextual variables. Origin-destination migration matrices and populations at risk were derived from census micro-data based on the temporally harmonized geography described above. Digital boundaries to match this geography were created by using 1992 and 2002 census municipal and census district digital boundaries obtained from the INE and the United Nations' Latin American and Caribbean Demographic Centre (UN-CELADE). The data set on regional attributes was built by drawing on information from the three censuses and external sources – including the Foreign Investment Committee (FIC), the Ministry of Housing and Urban Development (MINVU), the Chilean Chamber of the Construction Industry (CCHC), Dresdner, Sanhueza (2009) and Prado et al. (2002).

versus collected census data (Bianchini et al. 2013). A recommendation has been made to not use these data (Bravo et al. 2013).



Note: For a list of personal characteristics available in the CHIM database, see https://s3.amazonaws.-com/geoda/data/CHIM.zip.

Figure 2: Composite of aggregate data files in the CHIM database

2 Description of the resource

This section provides a description of the aggregate data sets from the CHIM database which are made available through this publication. Open access to census micro-data is not possible for confidentiality issues but customized micro-data requests can be made on http://www.franciscorowe.com. Four sets of data files are made available: (1) origin-destination migration matrices, (2) populations at risk, (3) digital boundaries, and (4) regional contextual variables. The content of each of these files are now described:

Origin-destination migration matrices This set of data files contains the output of 3,211 origin-destination migration matrices derived by cross-tabulating information on individuals' usual place of residence at the census date and five years earlier. The data cover three five-year intervals: 1977-1982, 1987-1992 and 1997-2002, and are stored in long format with each row corresponding to an origin-to-destination pair. They include the diagonal elements of the matrices, showing the number of people staying in the sample geographical area. The data are disaggregated by five-year age groups, 2 gender categories, 5 educational attainment levels, 5 employment classes, 3 labor force statuses, 11 industry sectors, and 10 occupational groups. The data are available at the TM scale and can be aggregated to provincial and regional geographical scales by using the respective identifiers in the metadata file.

Populations at risk This set of data files contains place-specific populations which correspond to the row sum of origin-destination migration matrices and indicate the population at the start of each census interval.

Digital boundary This is a folder which consists of a set of Geographic Information System (GIS) files, providing digital boundaries for three spatial scales: TMs, Provinces and Regions.

Regional contextual variables This data file contains, as indicated in Figure 1, a set of labor market, housing market, economic liberalization, and regional composition and structure variables.

These data have already been deployed to examine the determinants of long-distance commuting (Rowe 2014, Rowe, Bell forthcoming) and to build functional labor market areas for the analysis of labor migration (Casado-Díaz et al. 2017). However, they offer an opportunity for many other existing applications and future work will update the CHIM database to include data from the 2017 census.

3 Potential applications

Internal migration is widely recognized as an integral component of national development. As countries modernize and become globally connected, the intensity, forms, and patterns of population movement evolve, reshaping patterns of human settlement. Chile provides an ideal exemplar. Over the last four decades, Chile transitioned from a closed, centrally-planned economic system with a socialist government to a globally integrated, market-driven economy under military and democratic regimes. This has transformed Chile's space economy. Prior work has systematically traced the evolution of mobility in Chile and its connections to the socio-economic and political context, revealing a declining trend in the intensity and redistributing capacity of internal migration (Rowe 2013).

The described database offers a unique opportunity to understand the forces driving these declining trends in the context of a global decline in internal migration intensity. Identification of the causes underpinning the migration decline would provide empirical-based evidence to guide policy makers to ensure appropriate economic and social policy responses to the continuing decline in human mobility. This finding also has the potential to inform sub-national population and housing forecasting, and enable the formulation of polices that ensure that regions with economic opportunities can attract the required workers to enhance productivity and build resilient local economies.

The described database also offers an opportunity to examine long-term associations between internal migration and industry agglomeration patterns, and understand the effects of regional economic performance and policy changes on population movements across the Chilean urban hierarchy. Empirical evidence suggests that as urban systems mature, they develop a hierarchical network of primary, intermediate, and small cities (Kontuly, Geyer 2005). Cities appear to have progressed through cyclical concentration and deconcentration phases, starting with urbanization, followed by polarization reversal, counter-urbanization and re-urbanization. Differences in political and socio-economic conditions are noted to affect the timing of these processes, and each of these processes is characterized by particular patterns of migration, reflecting the spatial diffusion of urban development and economic growth. Urbanization is characterized by net in-migration into major urban areas, echoing the spatial concentration of development and economic growth. Polarization reversal is associated with deconcentration of firms and capital to peripheral areas of smaller population size due to diseconomies of scales in major urban centers. Counter-urbanization comprises net out-migration from urban centers, reflecting the inner city economic decline and mobility to rural locations; and re-urbanization involves gentrification and re-population of city centers through net in-migration. Organizing the TM into a functional urban framework, the CHIM database can be deployed to explore how these processes have played out in Chile and how they have been shaped by contextual changes.

References

Bianchini R, Feeney G, Singh R (2013) Report of the international commission on the 2012 population and housing census of Chile. Instituto nacional de estadísticas, Santiago, Chile

Bravo D, Larrañaga O, Millán I, Ruiz M, Zamorano F (2013) Informe final comisión externa revisora del censo 2012. Instituto nacional de estadísticas, Santiago, Chile

Casado-Díaz JM, Martínez-Bernabéu L, Rowe F (2017) An evolutionary approach to the delimitation of labour market areas: an empirical application for Chile. *Spatial Economic Analysis*: 1–25. CrossRef.

Dresdner J, Sanhueza C (2009) Estimación de series de salarios regionales. *Estudios de Economía* 36[2]: 243–261

Kontuly T, Geyer H (2005) Lessons learned from testing the differential urbanisation model. *Tijdschrift voor economische en sociale geografie* 94: 124–128. CrossRef.

Prado M, Gatica R, Letelier J (2002) Las exportaciones regionales de Chile en la perspectiva de la globalización: Período 1990-2000. Technical report MIDEPLAN, Santiago, Chile

Rowe F (2013) Spatial labour mobility in a transition economy: Migration and commuting in Chile. PhD thesis, school of geography, planning and environmental management, The University of Queensland, Brisbane, Australia. CrossRef.

Rowe F (2014) Micro and macro drivers of long-distance commuting in Chile: The role of spatial distribution of economic activities and population. paper presented at the 53rd congress of the Western Regional Science Association, San Diego, CA

Rowe F, Bell M (2013) Creating an integrated database for the analysis of spatial mobility in Chile. Working paper 02/2013, Queensland centre for population research, school of geography, planning and environmental management, The University of Queensland, Brisbane, Australia

Rowe F, Bell M (forthcoming) The drivers of long-distance commuting in Chile: The role of spatial distribution of economic activities and population. In: Poot J, Roskruge M (eds), *Regional science perspectives on population change and impacts in Asia and the Pacific*. Springer

UNESCO – The United Nations Educational Scientic and Cultural Organization (1997) International standard classication of education. ISCED 1997, 22nd May, 2012. http://www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 3, 2017, R7–R9
DOI: 10.18335/region.v4i3.199

journal homepage: region.ersa.org

Functional Labor Market Areas for Chile

Francisco Rowe[1], José Manuel Casado-Díaz [2], Lucas Martínez-Bernabéu[3]

[1] University of Liverpool, Liverpool, UK (email: F.Rowe-Gonzalez@liverpool.ac.uk)
[2] University of Alicante, Alicante, Spain (email: jmcasado@ua.es)
[3] University of Alicante, Alicante, Spain (email: lucas.martinez@ua.es)

Received: 21 July 2017/Accepted: 22 July 2017

Abstract. Administrative areas are arbitrarily designed and do not necessarily reflect the geographical patterns of socio-economic and labor market activity. Labor market areas (LMAs) are required to analyze spatial labor market activity and provide a framework to guide spatially-explicit employment policy development. This resource describes a data source of a set of recently created labor market areas for Chile.

1 Background

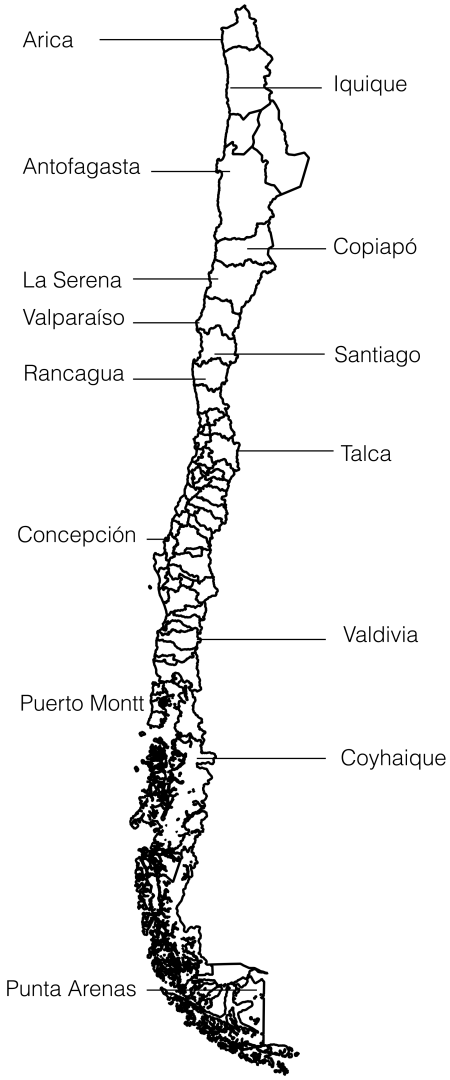
In 2009, a comprehensive review of Chile's territorial economic performance by the OECD highlighted underutilized potential in the use of regional assets and capacity to boost national economic productivity (OECD 2009). To redress these deficiencies, recommendations to transition to a territorial policy approach were emphasized and Casado-Díaz et al. (2017) developed a set of functional Labor Market Areas (LMAs) to assess spatial labor market activity, and to guide future employment policy development.

2 Description of the resource

The data set includes 62 LMAs, providing full coverage of the Chilean territory (Figure 1). The LMAs were delineated by means of the algorithm described in Casado-Díaz et al. (2017) which uses evolutionary computation to implement an optimization process. In this process the overall internal cohesion of LMAs is maximized subject to restrictions of minimum levels of self-containment and population size which are applied to each of the resulting LMAs. The method is an extension of the algorithm proposed by Martínez-Bernabeu et al. (2012). The regionalization process was performed using commuting data from the CHilean Internal Migration (CHIM) database (Rowe, Bell 2013). Data from the 1982, 1992, and 2002 Chilean Housing and Population census were appended to the set of LMAs to produce a geographic information database. The database contains information on the resident population by five-year age groups, sex, labor force status, industry sector, and occupation and is available in a shapefile format. Together with this database, a correspondence file for the matching municipality and labor market areas is made available.

3 Potential applications

This database provides a unique opportunity to better understand the economic geography of Chile over a period of twenty years, between 1982 and 2002. This was a period



Notes: Labels for regional capital cities are displayed

Figure 1: Functional labor market areas for Chile

of significant economic transformation in Chile paralleling its transition to a globally integrated, market-driven system from a closed, centrally-planned economy. Drawing on evolutionary economic geography (Boschma, Frenken 2006) and economic complexity theory (Hidalgo, Hausmann 2009), the resource described above offers the possibility to enhance our understanding of the robustness, plasticity, niche construction, and evolvability of regional economic systems, i.e., how regional economic systems maintain certain functionalities, adapt, shape their immediate environments, and evolve in response to environmental disturbances (Martin, Sunley 2015).

The proposed resource can also be employed to trace the evolution of the interaction between labor demand and supply components in Chilean labor markets over three census periods, extending prior analysis (Rowe 2013, 2014, Rowe, Bell forthcoming). A labor market account (Rowe 2013) and a panel vector autoregression approach (Blanchard et al. 1992) can be implemented to identify the major routes of regional labor market adjustment, the way regional labor markets adjust to shifts in employment, and relative importance of geographic mobility processes, such as internal migration and commuting, play in regional labor markets. Such analyses have the potential to guide regional labor market policy development by providing evidence on the responsiveness of these processes to job creation and its potential leakages to non-local workers.

Acknowledgement

This work was supported by the Spanish Ministry of Economy and Competitiveness [grant numbers CSO2011-29943-C03-02 and CSO2014-55780-C3-2-P, National R&D&i Plan] and the Chilean National Commission for Scientific and Technological Research through its human capital development program.

References

Blanchard OJ, Katz LF, Hall RE, Eichengreen B (1992) Regional evolutions. *Brookings Papers on Economic Activity* 1[1]: 1–75. CrossRef.

Boschma R, Frenken K (2006) Why is economic geography not an evolutionary science? towards an evolutionary economic geography. *Journal of Economic Geography* 6[3]: 273–302. CrossRef.

Casado-Díaz JM, Martínez-Bernabéu L, Rowe F (2017) An evolutionary approach to the delimitation of labour market areas: an empirical application for Chile. *Spatial Economic Analysis*: 1–25. CrossRef.

Hidalgo C, Hausmann R (2009) The building blocks of economic complexity. *Proceedings of the National Academy of Sciences* 106[26]: 10570–10575. CrossRef.

Martin R, Sunley P (2015) On the notion of regional economic resilience: conceptualization and explanation. *Journal of Economic Geography* 15[1]: 1–42. CrossRef.

Martínez-Bernabeu L, Flórez-Revuelta F, Casado-Díaz J (2012) Grouping genetic operators for the delineation of functional areas based on spatial interaction. *Expert Systems with Applications* 39[8]: 6754–6766. CrossRef.

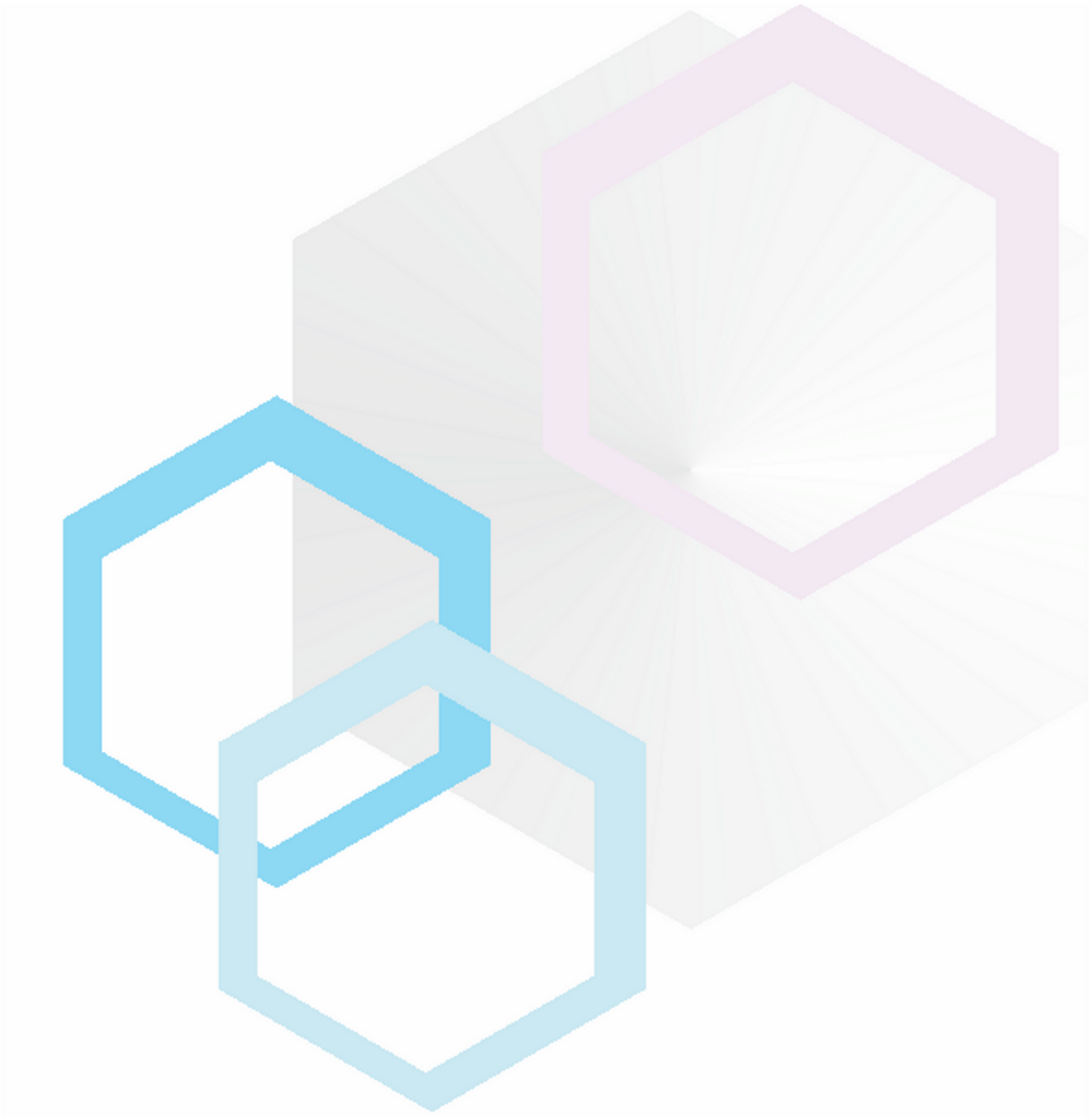
OECD – Organisation for Economic Cooperation and Development (2009) Territorial reviews: Chile 2009. OECD publishing, Paris, France

Rowe F (2013) Spatial labour mobility in a transition economy: Migration and commuting in Chile. PhD thesis, school of geography, planning and environmental management, The University of Queensland, Brisbane, Australia. CrossRef.

Rowe F (2014) Micro and macro drivers of long-distance commuting in Chile: The role of spatial distribution of economic activities and population. paper presented at the 53rd congress of the Western Regional Science Association, San Diego, CA

Rowe F, Bell M (2013) Creating an integrated database for the analysis of spatial mobility in Chile. Working paper 02/2013, Queensland centre for population research, school of geography, planning and environmental management, The University of Queensland, Brisbane, Australia

Rowe F, Bell M (forthcoming) The drivers of long-distance commuting in Chile: The role of spatial distribution of economic activities and population. In: Poot J, Roskruge M (eds), *Regional science perspectives on population change and impacts in Asia and the Pacific*. Springer



Funded by

ersa

WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

FWF