Volume 4, Number 1, 2017

REGION

The Journal of ERSA
Powered by WU

Table of Contents

Articles

Funded by ersa WU FWF

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

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

Articles

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 1, 2017, 1–16
DOI: 10.18335/region.v4i1.95

journal homepage: region.ersa.org

Does the increase in house prices influence the creation of business startups? The case of Sweden

Björn Berggren[1], **Andreas Fili**[2], **Mats Håkan Wilhelmsson**[3]

[1] Royal Institute of Technology, Stockholm, Sweden (email: bjorn.berggren@abe.kth.se)
[2] Royal Institute of Technology, Stockholm, Sweden (email: andreas.fili@abe.kth.se)
[3] Royal Institute of Technology, Stockholm, Sweden (email: mats.wilhelmsson@abe.kth.se)

Received: 22 September 2015/Accepted: 22 November 2016

Abstract. Entrepreneurs are at the core of economic development in that they start new businesses or make existing firms grow. To fulfill this important role, entrepreneurs need access to financing. Owing to information asymmetry and the relatively high risk associated with business startups, many financiers shy away from engaging in relationships with firms during the early stages of their development. Based on the existing body of knowledge on the financing of entrepreneurship, we know that insider finance is of paramount importance in the early stages of firms' development. We expand this knowledge base by analyzing the influence of house prices on business startups across municipalities in Sweden. In our analysis, we include data from all municipalities in Sweden. Our data on house prices and control variables are collected in period one, and our data on the frequency of startups are collected in period two. We find that rising house prices in a municipality lead to a higher frequency of startups. In our spatial Durbin model, we find that a 1% increase in house prices leads to around 0.15% increase in startups. Our findings are in line with the limited international research that has been previously conducted, and therefore, our study might make a small but vital addition to this growing body of knowledge within the area of entrepreneurship and regional development.

JEL classification: R11, R31, M13

Key words: Business startups, entrepreneurship, financing, house prices, mortgages

1 Introduction

In the late 1970s, scholars concluded that small and medium-sized enterprises (SMEs) create the majority of new jobs in the U.S. economy (Birch 1979). These findings spurred great interest in research on the employment contribution of SMEs worldwide, with scholars concluding that SMEs contribute to 70% to 90% of all new jobs that are created (Davidsson et al. 1995, Armington, Acs 2002, Santarelli, Tran 2012). Previous research has shown that in addition to playing a vital role in the creation of employment opportunity, startups and SMEs are involved in creating industrial renewal, export income, and innovation (Halilem et al. 2012, Agostini et al. 2015, Love, Roper 2015), as well as acting as a dynamic influence to lagging areas (Keeble 1997, Gordon, McCann 2005, Doh, Kim 2014)(Keeble 1997, Gordon, McCann 2005, Doh, Kim 2014). Therefore, governments in most nations have developed different types of programs to support SMEs and startups (Bateman 2000, Perren, Jennings 2005).

To fulfill the role of creators of new employment, export income, and innovations, new firms must secure access to financial resources (Harding, Cowling 2006, Atherton 2012). Previous research has shown that the most important source of finance for newly started firms is insider finance (Cassar 2004, Gregory et al. 2005, Robb, Robinson 2014). Insider finance includes the personal funds of the founder: personal savings, home mortgage, and credit cards (Storey, Greene 2013). This implies that a booming housing market would enhance the ability of entrepreneurs to increase their home mortgage, or at least enhance their ability to take on new loans by virtue of the increase in collateral they can offer the bank, and, thus, finance their newly started firms – a causality that has been established by Jin et al. (2012).

In this paper, we build on the findings of Jin et al. (2012) and use Sweden as our empirical case. More specifically, we attempt to identify and estimate an empirical model of the relationship between business startups and house prices in all 290 Swedish municipalities.

The remainder of the paper is organized as follows: In section 2, we present some key findings from previous studies on the financing of entrepreneurial ventures and the relationship between house prices and business startups. Next, we present the empirical model in Section 3 and describe our data in Section 4. We then present our results in Section 5 and conclude the paper with a discussion of our findings in Section 6.

2 Theoretical points of departure

Financing entrepreneurship and business startups have received a great deal of attention from researchers and policymakers for over 100 years (see MacMillan 1931). According to conventional wisdom, small and medium-sized firms have problems accessing finance at reasonable terms (see Storey, Greene 2013). Whether SMEs are, in fact, subject to credit rationing is a question that has been asked in numerous research studies but is one that remains unanswered. A consensus, however, has been reached regarding the dependency of entrepreneurial ventures on insider finance in their earlier stages of development as well as regarding the higher degree of financial constraints experienced by SMEs in comparison to larger organizations (Cassar 2004, Revest, Sapio 2012). As a consequence, there has been considerable research interest in the financing of new firms. Three major theories have been used to illustrate financing patterns: life-cycle theory, pecking order theory, and agency theory (Berger, Udell 1998, Johnsen, McMahon 2005). The first two theories focus on the perspective of the firm that receives financing, whereas agency theory takes the perspective of the investor that provides financing.

2.1 Different perspectives on the financing of entrepreneurial ventures

Life-cycle theory suggests that firms, depending on what stage of development they have reached, follow similar financing patterns (Weston, Brigham 1981). Research has revealed that firms in certain stages of development seek certain types of financing and that firms have similar financing needs and financing behavior – no matter the cultural differences across countries (cf. Psillaki, Daskalakis 2009). Studies from all over the world – including Europe (Psillaki, Daskalakis 2009), China (Newman et al. 2012), and Africa (Abor, Biekpe 2009) – have provided support for this perspective.

Pecking order theory is concerned with explaining why firms do not always prefer the source of financing with the lowest interest rate. Research has shown that there seems to be a stable preference order – a pecking order – whereby different sources of financing are ranked (Donaldson 1984, Myers 1984). In essence, the theory states that there is a general mistrust of outsiders: the more a firm is likely to lose control to external financiers, the less likely it is to submit to that type of financing. Internally generated funds are preferred to bank loans, which in turn are preferred to new equity. Although initial studies were conducted in large companies, several studies have shown that the pecking order framework is a fruitful approach to studying financial decision-making in small firms as well (Vanacker, Manigart 2010, Degryse et al. 2012, Alon, Rottig 2013) and that SMEs follow a financial pecking order when they seek external financing (Berggren et al. 2000). Like life-cycle theory, pecking order theory is a well-established theory, and numerous studies have found support for it (cf. Davidsson et al. 2009, Mac an Bhaird, Lucey 2011).

Whereas both life-cycle theory and pecking order theory focus on the perspective of the firm, agency theory takes the perspective of the investor. Agency theory models entrepreneurial finance in terms of contracts between a principal and an agent, where the goals of the two parties diverge (Jensen, Meckling 1976). Thus, as soon as the investor has supplied funding, the firm will try to use those resources for personal gains instead of for the benefit of the investor. The fact that these goals differ implies that once the funding has been supplied, the investor needs to ensure – through monitoring and control – that the funds are used properly (Ross 1973).

2.2 Implications of the financial perspectives for entrepreneurial finance

The major explanatory construct in pecking order theory is the notion of control aversion. The pecking order theory predicts that banks will be preferred to new shareholders. However, this notion is also quite clearly linked to life-cycle theory, in that control aversion is especially prevalent among young firms: overcoming control aversion is partly a matter of reaching maturity in dealing with business associates, financiers, and presumptive owners. Because control aversion is prevalent among young firms, they will often act in accordance with pecking order theory by contacting banks, rather than new owners (Howorth 2001, Paul et al. 2007).

The bank, in line with agency theory, will also demand some measures of control, primarily through collateral provided personally by the founder(s). Thus, the very foundational assumption of agency theory leads to a strong focus on control, which was initially the reason for the firm's decision not to seek other sources of financing. The lazy bank hypothesis states that collateral is not an effective measure against bankruptcy, but merely an easy way of handling SMEs (Manove et al. 2001). Still, as the bank retains the right to cancel a loan at any time, collateral, performance data and legitimacy represent significant obstacles to new enterprises (Bracke et al. 2013, De Clercq et al. 2013, Ramlall 2014). Because of the pivotal role of collateral, the size of the collateral also proves to be important. Being able to offer a large portion of valuable collateral to the bank should mean more access to financing.

Most young firms have nominal collateral and credibility to offer the bank: they are not as transparent as older, larger firms (Robb, Robinson 2014). Moreover, service firms have fewer tangible assets (machinery and inventory) to offer as collateral for loans than do manufacturing firms. Instead, collateral is found in the personal property of the founder and, therefore, main sources of collateral are the houses of the founders' (Chaney et al. 2012).

2.3 A model of higher house prices leading to an increase in business starts

Previous research shows that collateral is important for entrepreneurial activity. By using variations in house prices, Schmalz et al. (2013) provide evidence that in regions with house price appreciation homeowners are more likely to start a business; and the firms started by homeowners are larger than those started by renters. That is, collateral matters. Furthermore, Adelino et al. (2015) show, based on aggregated county data for the period 1998-2010, that regions with larger rises in house prices experienced stronger growth in employment in small firms, especially in industries with a limited need for capital.

The model is explained theoretically in the following way. First, higher house prices mean that nascent entrepreneurs have more collateral to offer to the bank when they apply for a loan to start the business (cf. Bernanke et al. 1999, Greenspan, Kennedy 2008, Jin et al. 2012, Bracke et al. 2013). This implies that banks can grant more loans to small businesses based on the fact that there is more private collateral on the part of the owner-founder of the firm. The model is part of an argument that claims that entrepreneurs suffer from a lack of funding (cf. MacMillan 1931, Stiglitz, Weiss 1981) and that sudden financial gains – "windfall gains" – have a positive effect on the number of business starts (Schäfer et al. 2011).

A number of papers lend support to this model. In their examination of the number of value added tax (VAT) registrations in the United Kingdom as well as other kinds of aggregate data, Black et al. (1996) found evidence that collateral availability does have a strong influence on firm formation and dissolution. By contrast, Hurst, Lusardi (2004) found no support for this hypothesis except in the very richest segment of households. Thus, they conclude that there is no general link between collateral and business starts. Robson (1996) compares his results explicitly with those of Black et al. (1996) from the same year and also in the United Kingdom. Although Robson (1996) finds no support for a link between high house prices and an increase in business starts, he states that housing wealth does appear to have a positive effect on entrepreneurship, in that it helps reduce the regional rate of deregistration from VAT.

This implies that an increase in house prices would equal an increase in the pool of liquidity available to entrepreneurs, which translates into a larger number of business starts in a certain region. According to this logic, a causal link may exist between a rise in house prices and the number of businesses started in a region as a result of greater access to collateral. Thus, our null hypothesis is formulated as follows:

> There is a positive relationship between rising house prices and the number of business starts in a region.

3 Empirical model

To test our hypothesis that housing market conditions play an important role in explaining new firm formation, we estimate a model where the number of new firms per capita for the period 2007-2014 across Sweden is related to a number of determinants. Of main interest here is the relationship between new firm formation and housing price growth.

Two major problems may arise in this type of regional model. The first problem involves the issue of endogenous determinants. Because house price appreciation facilitates lending and, ultimately, increases entrepreneurial activity and economic growth, higher economic activity might result in more lending and higher house prices. That is to say, the relationship between house price appreciation and formation of new business startups is bidirectional (Iacoviello 2005, Adelino et al. 2015). The second problem involves the issue of spatial dependence. Both problems are connected to the question of how to interpret the estimated relationships as causality and not merely as correlations. That is, the empirical challenge is to identify the causal direction of the house price growth effect.

We avoid these potential problems by estimating a model using determinants in the preceding period (2007) when explaining the variation in firm formation in the subsequent period (2007-2014). In contrast to Binet, Facchini (2015), we estimate three types of spatial autoregressive models, a spatial lagged model and a spatial error model, as a means to control spatial dependence. We also estimate a spatial Durbin model in order to analyze if it could be simplified to a spatial error or lag model. Five different spatial weight matrices are tested. Two nearest neighbor, two inverse distance based and, finally, a spatial contiguity matrix. The distance is estimated using the centroid coordinates of the labor market.

Our approach has been used recently by Andersson et al. (2014), and earlier by Armington, Acs (2002). Our identification strategy is to relate the change in the number of startups in a region with the change and level of house prices in the previous period. That is, lagged house prices can have an effect on startups in a later period, but it is unlikely that startups in the future have an effect on house prices in a previous period (cf. Balasubramanyan, Coulson 2013).

New firm formation per capita varies to a great extent across municipalities in Sweden for the period 2007-2014. We use five different types of determinants to explain this variation, namely, (a) establishment structure, (b) labor market conditions, (c) human capital, (d) income, and (e) housing market conditions.

We use establishment size, measured as 2007 employment divided by the number of establishments in 2007, as a proxy for the structure of the industry in the region (see Armington, Acs 2002). The expected relationship is negative, as larger average establishment size should be negatively related to regional startups rates (Armington, Acs 2002).

Different measures are used here to measure labor market conditions. Both measures, number of employees in 2007 and unemployment rate in 2007, are included in the model. Number of employees measures the size of the market and is used instead of population (Binet, Facchini 2015) or number of households (Adelino et al. 2015) as a measure of the market size. It is a measure that captures some of the agglomeration effects observed in the literature (see Acs et al. 1994). One of the most important determinants used in previous research is the unemployment rate (see Armington, Acs 2002) as it was suggested that unemployed workers were more likely to start new firms. The unemployment rate is also used in more recent research (Adelino et al. 2015, Binet, Facchini 2015). The third labor market indicator we use is the self-employment rate in 2007. This measure of entrepreneurial culture (Johannisson 1984), has also been used in research by Armington, Acs (2002).

We also include a variable measuring the percentage of the population that was born outside of Sweden in 2007. Lee et al. (2004) used this diversity index, which they labeled the "Melting Pot Index." The hypothesis is that a positive relationship exists between this index and new firm formation. The argument is that "immigrants lack skills, resources, and networks" and, therefore, tend to be self-employed and to start new companies to a greater extent than nonimmigrants. Migration data were also used by Adelino et al. (2015).

As a measure of human capital, we use the rate of university degree completion in the population in 2007. This measure is a proxy for the level of skill and knowledge in the regional economy (see Armington, Acs 2002, and, more recently, Adelino et al. 2015, Binet, Facchini 2015). The relationship between the rate of university degree completion in the population and new firm formation is expected to be positive.

Income is measured as the average annual regional income level in 2007 and is hypothesized to have a positive relationship with new firm formation in the subsequent 2007-2014 period. Based on the argument of Binet, Facchini (2015) that a high regional income level broadens the market size and, therefore, increases the number of opportunities for new firms, we would expect to observe more business startups in regions with higher regional income levels. We have also tested the change in annual income, but the empirical results suggest that it is not related to the number of new firms per capita. The level of income seems to be more important.

The housing market is included with two different measures. The first variable measures the annual growth in the seven years preceding 2007, that is, 2000-2006. The same type of measure is used by Schmalz et al. (2013) even though they use individual data. We also include a measure indicating whether the regional house price level in 2007 is above the average house price level.

4 Data

We use data on startups in Sweden from the period 2007 to 2014. The data are aggregated and based on all 290 municipalities in Sweden. The dependent variable is the change in the number of startups for the period 2007-2014. The independent variables all measured in 2007 are: human capital measured as the proportion with a university degree, income, employment, unemployment, and accessibility in the municipality as well as the change in house prices in the seven years preceding 2007 (2001-2007) and the house price level in 2007. Some descriptive statistics of the data are shown in Table 1.

In Table 1 we see that the average number of new firms per capita for the period is 0.05 (standard deviation: 0.01), which is equal to 5 new firms per 100 inhabitants. Almost 23% of the population has a university degree; the variation across the labor market is, however, substantial. House price appreciation is measured for the period 2001-2007. The average house price change is positive and equal to almost 0.7% (standard deviation: 0.3).

Table 1: Descriptive statistics of the data

Variable	**Abbre-viation**	**Period**	**Average**	**Standard deviation**
New firms per capita	New	2007–2014	0.0505	0.0142
University degree (%)	Univ	2007	0.229	0.078
High house prices (%)	High-Hp	2007	0.33	0.47
Change house price (%)	Dhp	2001–2007	0.6926	0.3216
Self-employment (%)	Self	2007	9.7391	2.4726
Employment	Emp	2007	15,203	30,722
Establishment size	Estab	2007	8.4618	1.7561
Income (SEK 000)	Inc	2007	210.426	18.208
Unemployment (%)	Unemp	2007	6.6398	1.9628
Immigration outside EU/EFTA (%)	Inm	2007	4.6474	3.3232
Stockholm (%)	Sthlm	2007	0.1211	0.3268

Source: EFTA, European Free Trade Association; EU, European Union.

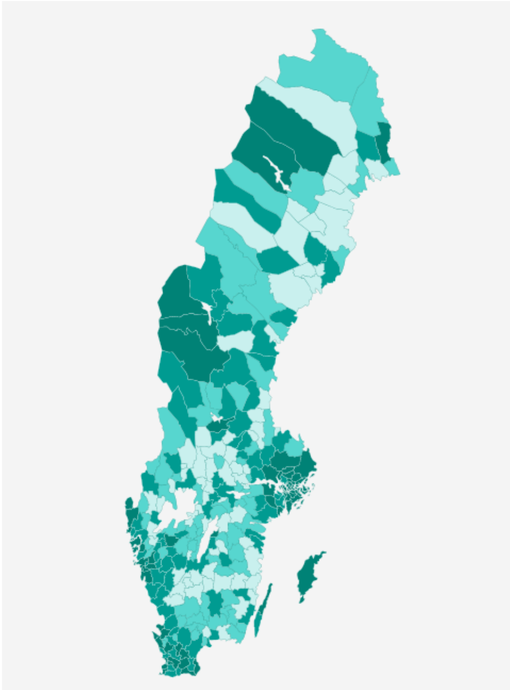
Table 2: Correlation matrix

	New	Univ	High-hp	Dhp	Self	Emp	Estab	Inc	Unemp	Inm
New	1									
Univ	0.63*	1								
High-Hp	0.71*	0.74*	1							
Dhp	0.62*	0.51*	0.59*	1						
Self	0.28*	-0.23*	-0.12*	0.03	1					
Emp	0.49*	0.74*	0.61*	0.48*	-0.42*	1				
Estab	-0.11	0.34*	0.26*	0.14*	-0.85*	0.45*	1			
Inc	0.34*	0.66*	0.66*	0.62*	-0.42*	0.57*	0.63*	1		
Unemp	-0.32*	-0.26*	-0.35*	-0.37*	-0.17*	-0.16*	-0.09*	-0.50*	1	
Inm	0.22*	0.38*	0.41*	0.29*	-0.47*	0.30*	0.47*	0.39*	-0.25*	1

Note: * Statistical significance on a 5% level.

Around one third of households in the labor markets have house prices that are higher than the average house prices. The average size of the labor markets, measured as the number of employees, is only 15,000 persons but the variation is considerable (standard deviation: 30,000 persons). The average unemployment rate is 6.6% with a variation of 2%. We measure the business set up in terms of establishment size. The average number of employees per establishment is equal to 8 persons (standard deviation: 1.7 persons). The entrepreneurial climate is measured with the self-employed variable. The average rate of self-employed is 9.7% with a variation of 2.5%. The number of immigrants as a percentage of the population is 4.6%, but the variation is substantial across the labor markets. Around 12% of the population lives in the labor market of Stockholm (the capital of Sweden). The correlation coefficients are shown in Table 2.

The correlations between the dependent variable (New) and the independent variables are strong in most cases. The highest correlation among the dependent variable and an independent variable is between new firms and high house prices, indicating the importance of house prices as a channel of financing for startups. However, we can also observe high correlations between high house prices and high proportion with a university degree and between high house prices and high levels of income, with university degree and income being positively correlated with startups. The house price appreciation for 2001-2007 is positively correlated with new firm formation for 2007-2014. We also notice that the variable 'employees per establishment' and the self-employed variable are highly negatively correlated, indicating that it can be difficult to differentiate the effects in the empirical model. All correlations are statistically significant on a 5% level. In Section 5, we present the results from our empirical model in which we relate new firm formation



Source: Arena for Growth (2015)

Figure 1: Average number of startups per 1000 inhabitants in Sweden 2011-2014

per capita to all the determinants presented here.

4.1 Entrepreneurship in Sweden

Before we present and analyze the findings of our empirical model, we will present some of the key characteristics of startups within the Swedish economy. The number of startups in Sweden has increased over the past 20 years. Between 1994 and 2003, there were 34,000-39,000 startups per year, but in 2014 that number had increased to more than 71,000 startups per year (Statistics Sweden 2015). Two of the major reasons for the increased number of startups are simplified rules for incorporating a business and reduced capital requirements for limited liability companies. The number of bankruptcies has been relatively stable over the past ten years and in 2014 6,000 Swedish firms filed for bankruptcy (Statistics Sweden 2015).

In comparison with other European countries, the number of nascent entrepreneurs in Sweden is relatively low (GEM 2012). One reason for the relatively low levels of startups in the Swedish economy, in comparison with other nations, is lack of entrepreneurial spirit owing to a tradition of large enterprises within the Swedish economy. Instead it seems as though most entrepreneurial skills are distributed among established firms, a phenomenon that has been labeled intrapreneurship (GEM 2012). Even though relatively few firms are started each year in Sweden, the firms that do start have a higher survival rate than firms started in comparable countries (Andersson, Klepper 2013). Regarding industries, most startups are within the retailing and services industries. In 2014, more than 80% of all new firms were started in these two industries.

In Sweden, there are relatively large regional differences in startup frequency, see Figure 1. Among the hotspots for startups, as well as being the most dynamic regions, are the three metropolitan areas of Stockholm, Gothenburg and Malmo. We can also find some examples of municipalities outside these regions with relatively high frequencies of startups. Among these are municipalities on the border to Norway, as well as various regional centers where universities, hospitals, and other governmental agencies and institutions are located.

5 Empirical results

Our main hypothesis is that the growth in house prices play a major role in explaining subsequent variation in new firms per capita. However, we also test the hypothesis that business set up, labor market conditions, human capital, and entrepreneurial climate play equally important roles when it comes to new firm formation. We test these hypotheses using a regression model where the dependent variable is the number of new firms per capita for the period 2007-2014. The independent variables used are those discussed in the previous section.

Tables 3-5 show the results from the ordinary least squares (OLS) estimation as well as from three types of spatial econometric models: the SAR-model (spatial autoregressive model), the SEM-model (spatial error model), and the spatial Durbin model (SDM).

For model comparison and selection of weight matrix we are following a specific-to-general test procedure proposed by Elhorst (2010). First we estimate an OLS and thereafter, using LM-tests (Anselin 1988, Anselin et al. 1996), testing for spatial dependency. If the LM-tests on OLS-residuals are significant, then SDM is estimated. If these LM-tests suggest that the SEM is the best spatial model, the log likelihood ratio test (LR-test) is used in order to test the convenience of SDM against SEM and SAR.

The next step is to select the spatial weight matrix. We are using three different types of weight matrices: nearest neighbor-based, distance-based and contiguity-based spatial matrix (discussed in for example Chasco 2013). All of them have been used in the spatial econometric literature (Elhorst 2010). Our selection of weight matrix is based on mobility pattern between municipalities. For example, contiguity and inverse distance have been used in Mendiola et al. (2015). We are using 2 and 10 nearest neighbor and 50 and 100 kilometers cut-off for the inverse distance matrix. The choice is somewhat arbitrary[1]. However, as Elhorst (2010) says, the wrong choice of the spatial weight matrix can distort the estimates, but "the probability that this really happens is small if spatial dependence is strong" (Elhorst 2010). LeSage, Pace (2014) call the belief that the estimates are sensitive for the choice of spatial weight matrix "the biggest myth in spatial econometrics". LeSage, Pace (2009), Stakhovych, Bijmolt (2009), and Halleck Vega, Elhorst (2013) are all in favor of using goodness-of-fit measures to discriminate among different spatial matrix specifications as there are no clear theoretical reasons for any specific form. We are using the most widely used log-likelihood value in order to differentiate between spatial weight matrices (Elhorst 2010). In order to test for the robustness of our coefficient estimates, we are analyzing the coefficients in the final spatial model specification using all the different spatial weight matrices.

As stated earlier, we are estimating the spatial Durbin model in order to test the hypothesis if this more general specification can be simplified with a spatial error (SEM) and/or autoregressive (lag) model (SAR). Here we are using the LR-test.

Table 3 shows the result from OLS. We are also testing for normality (Jarque-Bera), heteroscedasticity (Breusch-Pagan/Cook-Weisberg) and multicollinearity (Variance-of-inflation, VIF).

Around 83% of the variation in the total number of startups per capita between 2007 and 2014 can be explained by business set up, labor market conditions, and entrepreneurial climate, as well as by income and house price appreciation. The R^2 value is considerably higher than that reported by Lee et al. (2004), for example, but is of the same magnitude as reported by Armington, Acs (2002). All coefficients have the expected signs and are statistically significant on a 95% level. The t ratios are calculated using White heteroscedasticity-robust standard deviations as the Breusch-Pagan/Cook-Weisberg test shows presence of heteroscedasticity. The Jarque-Bera test shows that residuals are normally distributed, that is, Maximum Likelihood (ML) is a suitable estimation method of the spatial error and spatial lag models. See also Figure 2 where the kernel density estimate is compared to the normal distribution. In the OLS model, we also present the variance of influence (VIF) so as to analyze potential multicollinearity issues. High correlation among the independent variables does not seem to create a problem of

[1]We have also tested different numbers of nearest neighbors and different cut-offs, but that did not change the results and the overall conclusion.

Table 3: Regression analysis (OLS-results)

	Model 1 (OLS)		
	Coefficient	**t value**	**VIF**
High house prices	0.1244	6.22	3.04
Change in house prices	0.1944	7.04	2.01
Human capital	0.1795	3.64	4.54
Establishments	-0.1809	-2.51	7.06
Self-employed	0.3136	5.71	6.94
Employment	0.0271	2.01	3.71
Income	0.4419	2.5	4.96
Unemployment	0.0978	3.08	2.28
In-migration	0.0217	1.44	2.25
Stockholm	0.1577	6.87	1.9
Constant	-6.0839	-5.97	
R^2	0.832		
R^2-adjusted	0.826		
Jarque-Bera test (Prob >z)	0.145		
Breusch-Pagan/Cook-Weisberg test (Prob >chi^2)	0.007		
Observations	289		

Notes: OLS, ordinary least squares; VIF, variance of influence. White heteroscedasticity-robust standard deviation.

Table 4: Diagnostic tests (statistics and p-values)

Tests	Nearest neighbor		Inverse-distance		Contiguity
	2	10	50 km cut-off	200 km cut-off	
Spatial Error:					
Moran's I	4.985	9.244	5.155	5.155	7.385
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LM	21.379	62.846	39.584	29.281	44.261
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Spatial lag:					
LM	25.306	39.813	44.176	9.776	33.167
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

multicollinearity as VIF values are below 10.

Testing for spatial dependence (Moran's I and the LM-tests with different spatial weight matrices) reveals that we do have a problem with spatial autocorrelation and/or spatial heteroscedasticity (Wilhelmsson 2002). Table 4 shows the results of the diagnostic tests for spatial dependence in OLS regression. Five different weight matrices are tested: two nearest neighbor (2 and 10), different inverse-distances based matrices with different cut-offs (50 and 200 kilometers) and one spatial contiguity based matrix.

At least two conclusions can be drawn from the diagnostic tests. First, spatial dependence is present. All LM-tests are significant which indicates the presence of spatial dependence. Second, the diagnostic tests are in favor of the spatial error model as the LM-test concerning the spatial error model has a higher value compared to the value concerning the spatial lag model.

We continue our analysis to estimate a spatial Durbin model with all the different spatial weight matrices. We do that in order to test if the more general spatial Durbin model is preferred compared to the spatial error model and the spatial lag model. The test is carried out with a LR-test. The test statistics (33.24 and 46.25) are all higher than the critical value (18.31), which indicates that the hypothesis can be rejected. That is

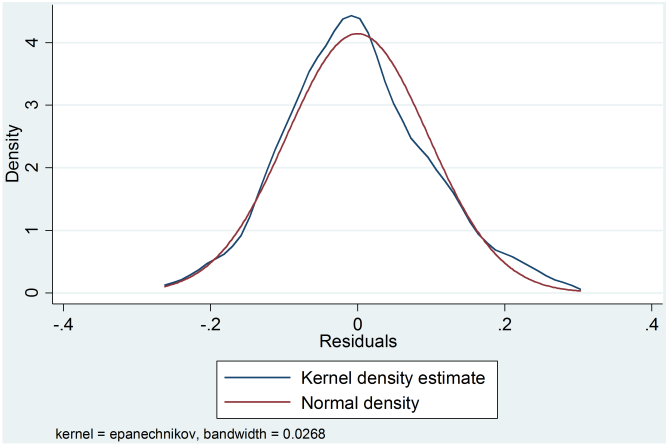


Figure 2: Kernel density estimate

to say, the spatial Durbin model describes the data best[2]. The results from the spatial Durbin model are presented in Table 5. The spatial weight matrix with the highest log likelihood is the contiguity matrix (see Table 6), and consequently, only the results from this specification are presented in Table 5 below. We have also estimated two diagnostic statistics concerning spatial autocorrelation in the residuals from the spatial Durbin model, namely LM-test and Moran's I. The Moran's I and LM-test is based on Anselin (2005) definition. Both of them show no indication of spatial dependency in the residuals.

By interpreting the coefficients we have the following results. If the proportion of the population with a university degree increases, the number of startups increases. That is, human capital is important as, for example, Armington, Acs (2002) and Lee et al. (2004) have shown. The same is true for an increase in the variables number of employees, income level, as well as unemployment rate. Our results are consistent with the findings of Armington, Acs (2002) and, more recently, Adelino et al. (2015) and Binet, Facchini (2015). The spillover effect (indirect effect) concerning human capital is negative indicating that lower human capital in neighboring municipalities is associated with fewer startups. However, the indirect effect concerning income and unemployment is positive, indicating spillover effect.

However, if the ratio between employment and establishment increases (the proxy for business set up), then the number of startups decreases. That is, regions with many SMEs are more likely to foster new startups. Our results concerning the Swedish market can confirm the results of Audretsch, Fritsch (1994) and those of Armington, Acs (2002), among others.

We can also observe that the number of self-employed persons in a labor market in 2007 leads to an increase in the number of startups per capita in the next seven years. Hence, the entrepreneurial climate seems to have an effect. The Melting Pot Index indicated that diversity is not significant in all models. Our results support the findings of Lee et al. (2004).

We can also notice a Stockholm effect. Being the capital, the Stockholm labor market fosters more startups per capita than the rest of Sweden. However, this effect alone cannot explain our findings. In fact, even if we exclude Stockholm, we can observe more or less the same results.

Housing market conditions represent the key determinant here. Two variables are used to measure housing prices. The first variable measures house price levels. It is a binary variable indicating whether the specific labor market has a price level that is above the national average house price level. The second variable measures house price growth. Both coefficients concerning housing market conditions are statistically significant and positive, indicating that a positive relationship exists between new firm formation and

[2]We have also tested SDM against SAR and the same conclusion can be drawn, that is the spatial Durbin model describes the data best.

Table 5: Spatial Durbin Model

	Coefficients	**Direct**	**Indirect effect**	**Total**
High house prices	0.0839	0.0885	0.0814	0.1699
	(4.25)	(4.40)	(1.33)	(2.55)
Change in house prices	0.1456	0.1591	0.2305	0.3896
	(5.30)	(5.77)	(2.84)	(4.59)
Human capital	0.2487	0.2242	-0.4373	-0.2131
	(6.40)	(5.92)	(-3.63)	(-1.66)
Establishments	-0.1683	-0.2882	-0.4229	-0.7111
	(-2.39)	(-3.98)	(-1.70)	(-2.68)
Self-employed	0.2111	0.1916	-0.3315	-0.1399
	(3.54)	(3.12)	(-1.51)	(-0.58)
Employment	0.0216	0.0248	0.0577	0.0825
	(2.02)	(2.32)	(1.67)	(2.22)
Income	0.2346	0.3191	1.6344	1.9535
	(1.55)	(2.09)	(3.44)	(3.76)
Unemployment	0.0708	0.0597	0.1805	0.2402
	(2.37)	(1.76)	(2.21)	(2.93)
In-migration	0.0226	0.0143	-0.031	-0.0164
	(1.61)	(1.00)	(-0.84)	(-0.42)
Stockholm	0.1891	0.1842	-0.0775	0.1068
	(3.91)	(3.87)	(-1.09)	(2.02)
Rho	0.422			
	(5.52)			
Constant	-7.7635			
	(-4.52)			
Log-likelihood	405.449			
R^2-adjusted	0.8484			
LR-test statistics (SDM vs SEM)	33.24			
LR-test statistics (SDM vs SAR)	46.25			
LM test for spatial autocorrelation (Prob >z)	0.8898			
Moran's I (Prob >z)	0.7492			

Note: *t*-values within brackets

both house price appreciation and house price level, respectively. Our finding supports the results of Adelino et al. (2015) and of Balasubramanyan, Coulson (2013).

Hence, the change in startups per capita for the period 2007-2014 can be explained by determinants that measure either the situation in 2007 or the change between 2001 and 2007. We argue that this is a causal relationship and not merely a correlation. If the change in house prices increases by 1%, the expected change in startups is around 0.15%. We also observe that the number of startups per capita is higher if the house prices are above average in 2007, indicating that both the level and the change in house prices are of importance. If we consider the direct impacts, we can observe that these are close to the spatial Durbin model coefficients, that is, the results indicate that the feedback effects are very small and of no economic significance. Spatial spillover measured by the indirect effect is positive and statistically significant. One interpretation could be that this positive spillover effect reflects how changes in house prices in all regions would impact startups in their own region. The spillover effect may be a result from expectations of future house prices. The indirect effect concerning the variable high house prices is, however, not statistically significant, indicating no regional spillover.

In order to test the robustness of our estimated coefficients concerning high house prices and the change in house prices we have also estimated the spatial Durbin model

Table 6: Spatial Durbin Model with different spatial weight matrices

Spatial weight matrix	**Change in house price**	**High house price**	**Log-likelihood**
Nearest neighbor 2	0.1556	0.0990	386.497
Nearest neighbor 10	0.1560	0.0817	399.016
Inverse distance 50	0.1465	0.0923	398.463
Inverse distance 200	0.2015	0.1268	390.913
Continuity	0.1456	0.0839	405.449

using different spatial weight matrices. The results are presented in Table 6. As observed, the difference in estimated parameters is very small, that is our results are robust when it comes to choice of spatial weight matrix. For example, the coefficient concerning the variable high house prices ranges from 0.0817 to 0.1268 and the coefficient concerning change in house price ranges from 0.1465 to 0.2015. Most of the estimates are of the same magnitude. The exception is the spatial weight matrix based on inverse distance with a cut-offs of 200 kilometers.

6 Conclusion and discussion

A number of interesting issues are highlighted in Section 5. For instance, most entrepreneurship theorists would agree that at the margin, there are nascent or potential entrepreneurs who lack access to finance. There also exist entrepreneurs at the margin who would remain self-employed longer if they had access to finance. However, assuming that people who suddenly receive money would generally use some or all of this money to start businesses would be an oversimplified view of the world. If potential entrepreneurs do not perceive an opportunity, or if they do not possess the unique capabilities necessary for exploiting a perceived opportunity, giving them money in itself is not enough (cf. Shane 2000).

In the present study, we have provided evidence that higher house prices – at an aggregate level – lead to an increase in business starts. A major contribution of our analysis lies in our modeling approach. We control for both endogeneity and spatial dependence of the entire population of Swedish municipalities. First, by separating observations in time, where the observation of house prices is in period n and the observation of entrepreneurship is in period n+1, we control for endogeneity and posit a causal relationship where higher house prices lead to an increase in entrepreneurial activity. Second, by employing a spatial Durbin model, we control for spatial dependence. We contribute to current theory by providing evidence in support of studies where house prices impact entrepreneurship. For policymakers, these results underline the paramount importance of the public sector's capacity for urban planning and the need for efficient processes in the institutional framework regulating the housing sector.

In the future, we intend to conduct a more specific analysis of how increases in house prices affect start-up frequencies in different sectors and regions. Today, the most common type of start-up in Sweden is a service firm with other firms as customers, but we expect to see regional differences in terms of the types of firms that are started.

One potential limitation of our paper is that it is based on an analysis of a single national case, in our case Sweden. Sweden has a bank-oriented economy (along with other countries such as Japan and Germany), whereas the United States and the United Kingdom are examples of market-oriented economies (cf. Mayer 1988). It is difficult to ascertain to what extent and in what way this orientation affects individual transactions and relations, and while there surely exist different traditions in different countries in terms of entrepreneurial activity, one could also make the opposite argument that some aspects of economic activity are increasingly global in nature and not very different between industrialized nations today. However, the difference in orientation has historically pervaded all economic activity. Therefore, future research in this field should endeavor to compare results between bank- and market-oriented economies.

References

Abor J, Biekpe N (2009) How do we explain the capital structure of smes in sub-Saharan Africa?: Evidence from Ghana. *Journal of Economic Studies* 36: 83–97. CrossRef.

Acs ZJ, Audretsch DB, Feldman MP (1994) R&D spillovers and recipient firm size. *Review of Economics and Statistics* 76: 336–340

Adelino M, Schoar A, Severino F (2015) House prices, collateral, and self-employment. *Journal of Financial Economics* 117: 288–306. CrossRef.

Agostini L, Caviggioli F, Filippini R, Nosella A (2015) Does patenting influence SME sales performance? A quantity and quality analysis of patents in Northern Italy. *European Journal of Innovation Management* 18: 238–257. CrossRef.

Alon I, Rottig D (2013) Entrepreneurship in emerging markets: New insights and directions for future research. *Thunderbird International Business Review* 55: 487–492. CrossRef.

Andersson M, Klepper S (2013) Characteristics and performance of new firms and spinoffs in sweden. *Industrial and Corporate Change* 22: 245–280. CrossRef.

Andersson R, Mandell S, Wilhelmsson M (2014) Så skapas attraktiva städer. Reforminstitutet, isbn 978-91-980600-8-9

Anselin L (1988) *Spatial Econometrics: Methods and Models*. Kluwer, Dordrecht. CrossRef.

Anselin L (2005) Exploring spatial data with geoDa™: A Workbook. Center for spatially integrated social science. revised version, march 6, 2005

Anselin L, Bera A, Florax R, Yoon MJ (1996) Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics* 26: 77–104. CrossRef.

Arena for Growth (2015) The new geography of sweden. Arena for growth, Stockholm

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

Atherton A (2012) Cases of start-up financing: An analysis of new venture capitalisation structures and patterns. *International Journal of Entrepreneurial Behavior & Research* 18: 28–47

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

Balasubramanyan L, Coulson E (2013) Do house prices impact business starts? *Journal of Housing Economics* 22: 36–44. CrossRef.

Bateman M (2000) Neo-liberalism, SME development and the role of business support centres in the transition economies of Central and Eastern Europe. *Small Business Economics* 14: 275–298

Berger AN, Udell GF (1998) The economics of small business finance: The roles of private equity and debt markets in the financial growth cycle. *Journal of Banking & Finance* 22: 613–673

Berggren B, Olofsson C, Silver L (2000) Control aversion and the search for external financing in swedish SMEs. *Small Business Economics* 15: 233–242

Bernanke BS, Gertler M, Gilchrist S (1999) The financial accelerator in a quantitative business cycle framework. In: Taylor J, Woodford M (eds), *Handbook of macroeconomics: vol. 1*. North Holland, Amsterdam. CrossRef.

Binet ME, Facchini F (2015) The factors determining firm start-ups in French regions and the heterogeneity of regional labor markets. *Annals of Regional Science* 54: 251–268. CrossRef.

Birch DGW (1979) The job generation process. MIT Program on Neighborhood and Regional Change 302

Black J, de Meza D, Jeffreys D (1996) House prices, the supply of collateral and the enterprise economy. *Economic Journal* 106: 60–75. CrossRef.

Bracke P, Hilber C, Silva O (2013) Homeownership and entrepreneurship: The role of commitment and mortgage debt. Institute for the Study of Labor (IZA) discussion paper no. 7417, IZA, Bonn

Cassar G (2004) The financing of business start-ups. *Journal of Business Venturing* 19: 261–283. CrossRef.

Chaney T, Sraer D, Thesmar D (2012) The collateral channel: How real estate shocks affect corporate investment. *American Economic Review* 102: 2381–2409. CrossRef.

Chasco C (2013) GeoDaSpace: A resource for teaching spatial regression models. *React@, Monográfico* 4: 119–144

Davidsson P, Lindmark L, Olofsson C (1995) Small firms, business dynamics and differential development of economic well-being. *Small Business Economics* 7: 301–315. CrossRef.

Davidsson P, Steffens P, Fitzsimmons J (2009) Growing profitable or growing from profits: Putting the horse in front of the cart? *Journal of Business Venturing* 24: 388–406. CrossRef.

De Clercq D, Lim D, CH O (2013) Individual-level resources and new business activity: The contingent role of institutional context. *Entrepreneurship Theory and Practice* 37: 303–330. CrossRef.

Degryse H, de Goeij P, Kappert P (2012) The impact of firm and industry characteristics on small firms' capital structure. *Small Business Economics* 38: 431–447. CrossRef.

Doh S, Kim B (2014) Government support for SME innovations in the regional industries: The case of government financial support program in South Korea. *Research Policy* 43: 1557–1569. CrossRef.

Donaldson G (1984) *Managing corporate wealth: The operation of a comprehensive financial goals system*. Praeger, New York

Elhorst JP (2010) Applied spatial econometrics. *Spatial Economic Analysis* 5: 9–28. CrossRef.

GEM (2012) Entrepreneurship in Sweden: National Report. Entrepreneurship Forum, Stockholm

Gordon IR, McCann P (2005) Innovation, agglomeration, and regional development. *Journal of Economic Geography* 5: 523–543. CrossRef.

Greenspan A, Kennedy J (2008) Sources and uses of equity extracted from homes. *Oxford Review of Economic Policy* 24: 120–144. CrossRef.

Gregory BT, Rutherford MW, Oswald S, Gardiner L (2005) An empirical investigation of the growth cycle theory of small firm financing. *Journal of Small Business Management* 43: 382–392. CrossRef.

Halilem N, Bertrand C, Cloutier JS, Landry R, Amara N (2012) The knowledge value chain as an SME innovation policy instrument framework: An analytical exploration of SMEs public innovation support in OECD countries. *International Journal of Technology Management* 58: 236–260. CrossRef.

Halleck Vega S, Elhorst J (2013) On spatial econometric models, spillover effects, and W. Working paper, Faculty of Economics and Business, University of Groningen

Harding R, Cowling M (2006) Assessing the scale of the equity gap. *Journal of Small Business and Enterprise Development* 13: 115–132. CrossRef.

Howorth CA (2001) Small firms' demand for finance: A research note. *International Small Business Journal* 19: 78–86. CrossRef.

Hurst E, Lusardi A (2004) Liquidity constraints, household wealth, and entrepreneurship. *Journal of Political Economy* 112: 319–347. CrossRef.

Iacoviello M (2005) House prices, borrowing constraints, and monetary policy in the business cycle. *American Economic Review* 95: 739–764. CrossRef.

Jensen MC, Meckling WH (1976) Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3: 305–360. CrossRef.

Jin Y, Leung CKY, Zeng Z (2012) Real estate, the external finance premium and business investment: A quantitative dynamic general equilibrium analysis. *Real Estate Economics* 40: 167–195. CrossRef.

Johannisson B (1984) A cultural perspective on small business – local business climate. *International Small Business Journal* 2: 32–43

Johnsen GJ, McMahon RGP (2005) Owner-manager gender, financial performance and business growth amongst SMEs from Australia's business longitudinal survey. *International Small Business Journal* 23: 115–142. CrossRef.

Keeble D (1997) Small firms, innovation and regional development in Britain in the 1990s. *Regional Studies* 31: 281–293. CrossRef.

Lee SY, Florida R, Acs Z (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*. Chapman & Hall/CRC, Boca Raton. CrossRef.

LeSage J, Pace RK (2014) The biggest myth in spatial econometrics. *Econometrics* 2: 217–249. CrossRef.

Love JH, Roper S (2015) SME innovation, exporting and growth: A review of existing evidence. *International Small Business Journal* 33: 28–48. CrossRef.

Mac an Bhaird C, Lucey B (2011) An empirical investigation of the financial growth lifecycle. *Journal of Small Business and Enterprise Development* 18: 715–731. CrossRef.

MacMillan HP (1931) Committee on finance and industry: report. Cmnd 4811, His Majesty's Stationery Office (hmso), London

Manove M, Padilla J, Pagano M (2001) Collateral versus project screening: A model of lazy banks. *RAND Journal of Economics* 32: 726–744. CrossRef.

Mayer C (1988) New issues in corporate finance. *European Economic Review* 32: 1167–1183. CrossRef.

Mendiola L, Gonzáles P, Cebollada A (2015) The relationship between urban development and the environmental impact mobility: A local case study. *Land Use Policy* 43: 119–128. CrossRef.

Myers SC (1984) The capital structure puzzle. *Journal of Finance* 39: 574–592. CrossRef.

Newman A, Gunessee S, Hilton B (2012) The applicability of financial theories of capital structure to the Chinese cultural context: A study of privately owned SMEs. *International Small Business Journal* 30: 65–83. CrossRef.

Paul S, Whittam G, Wyper J (2007) The pecking order hypothesis: Does it apply to start-up firms? *Journal of Small Business and Enterprise Development* 14: 8–21. CrossRef.

Perren L, Jennings PL (2005) Government discourses on entrepreneurship: Issues of legitimization, subjugation, and power. *Entrepreneurship Theory and Practice* 29: 173–184. CrossRef.

Psillaki M, Daskalakis N (2009) Are the determinants of capital structure country or firm specific? *Small Business Economics* 33: 319–333. CrossRef.

Ramlall I (2014) Is there a pecking order in the demand for financial services in Mauritius? *Journal of African Business* 15: 49–63. CrossRef.

Revest V, Sapio A (2012) Financing technology-based small firms in Europe: What do we know? *Small Business Economics* 39: 179–205. CrossRef.

Robb AM, Robinson DT (2014) The capital structure decisions of new firms. *Review of Financial Studies* 27: 153–179. CrossRef.

Robson MT (1996) Housing wealth, business creation and dissolution in the UK regions. *Small Business Economics* 8: 39–48. CrossRef.

Ross SA (1973) The economic theory of agency: The principal's problem. *American Economic Review* 63: 134–139

Santarelli E, Tran HT (2012) Growth of incumbent firms and entrepreneurship in Vietnam. *Growth and Change* 43: 638–666. CrossRef.

Schäfer D, Talavera O, Weir C (2011) Entrepreneurship, windfall gains and financial constraints: Evidence from Germany. *Economic Modelling* 28: 2174–2180. CrossRef.

Schmalz MC, Sraer DA, Thesmar D (2013) Housing collateral and entrepreneurship. Working paper 19680, National Bureau of Economic Research (NBER), Cambridge, MA. CrossRef.

Shane SA (2000) *A general theory of entrepreneurship: The individual-opportunity nexus*. Edward Elgar, Cheltenham

Stakhovych S, Bijmolt THA (2009) Specification of spatial models: A simulation study on weights matrices. *Papers in Regional Science* 88: 389–408. CrossRef.

Statistics Sweden (2015) Business activities. http://www.scb.se/en_/Finding-statistics/Statistics-by-subject-area/Business-activities/

Stiglitz JE, Weiss A (1981) Credit rationing in markets with imperfect information. *American Economic Review* 71: 393–410

Storey DJ, Greene FJ (2013) *Small business and entrepreneurship*. Pearson, Harlow

Vanacker TR, Manigart S (2010) Pecking order and debt capacity considerations for high-growth companies seeking financing. *Small Business Economics* 35: 53–69. CrossRef.

Weston JF, Brigham EF (1981) *Managerial finance* (7th edition ed.). Dryden Press, Hinsdale, IL

Wilhelmsson M (2002) Spatial models in real estate economics. *Housing, Theory and Society* 19: 92–101. CrossRef.

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 1, 2017, 17–30
DOI: 10.18335/region.v4i1.102
journal homepage: region.ersa.org

Analysis of Freight Trip Generation Model for Food and Beverage in Belo Horizonte (Brazil)*

Leise Kelli de Oliveira[1], Rodrigo Affonso de Albuquerque Nóbrega[2], Daniel Gonçalves Ebias[3], Bruno Gomes e Souza Corrêa[4]

[1] Universidade Federal de Minas Gerais, Belo Horizonte, Brazil (email: leise@etg.ufmg.br)
[2] Universidade Federal de Minas Gerais, Belo Horizonte, Brazil (email: raanobrega@ufmg.br)
[3] Centro Universitário Formiguense, Formiga, Brazil (email: daniel_ebias@outlook.com)
[4] Universidade Federal de Minas Gerais, Belo Horizonte, Brazil (email: bgesc@hotmail.com)

Received: 20 November 2015/Accepted: 22 December 2016

Abstract. Today, one of the main challenges faced in urban logistics is the distribution of goods. In Brazil, mid to large cities have experienced negative consequences of unplanned urban sprawl as well as the lack of adequate transport infrastructure. As for the logistic standpoint, the relationship between urban planning and transportation planning must be better exploited towards the promotion of quality of life, and economic and environmental sustainability. The attractiveness of urban activities that attract the movement of people and goods and other component elements of urban space should be investigated. The presence of bars and restaurants falls within this context and is therefore vital and responsible for a significant percentage of jobs and revenue in a city. This paper presents the results of a freight trip generation model developed for pubs and restaurants in Belo Horizonte (Brazil). The data for the freight trip generation model were obtained by survey. A structured questionnaire was designed to obtain information about goods, frequency, operational time, place of performance of the loading/unloading of goods, establishment size and the number of employees. In creating the proposed models, a simple linear regression was applied to correlate the following variables: (i) number of trips versus area of the establishment; (ii) number of trips versus number of employees; (iii) number of trips versus operation day of the establishment. With the results of the linear regression for trip generations, data interpolation was conducted based on the standard deviation of the results to define the sample classification bands. Finally, the resulting trip generation surface was analysed together with other geographic data such as demographic data, road network density and socioeconomic data. Findings indicate the importance of a mathematic-geographic model for trip generation as a feasible approach to support transportation planning & operation for urban goods distribution. Critical information such as the high concentration of pubs and restaurants in the same region can reinforce the vocation of the city for trading. However, an elevated number of freight vehicles to meet a high and growing demand becomes a problem especially in areas where the urban road network is not efficient (not properly designed and parking spaces not properly used). This study also highlights the need for an urban freight mobility plan and public policies, by offering sustainable alternatives for urban goods distribution,

*The support of the National Council for Scientific and Technological Development (CNPq) is acknowledged and appreciated. The authors thank the Municipality of Belo Horizonte for the database, which allowed the analysis of this study. The authors would also like to thank anonymous reviewers for the suggestions to improve this paper.

which improve the urban environment. By using geospatial analysis, the study delivered statistics data and maps to catch the attention of decision makers and transportation managers, therefore facilitating the discussion on transportation policies in the city of Belo Horizonte.

Key words: urban goods distribution, freight trip generation, trucking, pubs and restaurants, geospatial analysis, Brazil

1 Introduction

Nowadays, the distribution of urban goods is one of the main challenges faced by policy makers, transportation authorities and the private sector in Brazilian cities. These cities present mostly a pattern of non-organized growth, lack of planning, and especially a lack of adequate infrastructure. Other than these well known factors there is also an increasing rate of population density. According to the Brazilian Institute of Geography and Statistics (IBGE 2010), about 83.4% of Brazil's population lives in urban areas. Concomitantly to the lack of planning for the use of urban space, the increasing volume of vehicles is also aggravating the situation. The population growth in the city increases the demand for goods and services that potentiate the problems related to urban mobility and urban freight transport. According to Holguín-Veras et al. (2012), the relationship among land use and freight represents a central issue for adequately planning investments in infrastructure and land use policy. However, the current transportation planning process in Brazil does not estimate freight activity to assist decision makers, which is critical information to be considered when making infrastructure investment choices. This fact motivated the development of this study, which presents data and models to estimate the flow of urban freight vehicles.

In this context, we hypothesize in this investigation that a transportation model sensitive to geographic-based context will add quality and efficiency to computing freight trip generation by considering local characteristics. Thus, this paper's analysis of a freight trip generation model was computed based on a survey developed for pubs and restaurants in the city of Belo Horizonte, Brazil. We analyse pubs and restaurants due to this sector's importance in Belo Horizonte (area of study); it represents 42% of the number of deliveries in the central Area (Oliveira 2014, Oliveira, Guerra 2014). Besides that, what makes this study distinct is its characterization of the sector and the geographical analysis of the results, which intends to inform transport policies for the distribution of food and beverage in pubs and restaurants. The methodology proposed in this study intends to provide a guideline for planning of urban freight transport in order to obtain data for freight transport policies through field surveys (this type of analysis is still incipient in the Brazilian context).

The paper is structured as follows: Section 2 provides a brief literature review on urban freight trip generation models in the Brazilian context; Section 3 describes the study area; Section 4 describes the methodology and data; results are illustrated in Section 5; and Section 6 concludes and discusses future research.

2 Brief notes about urban freight trip generation models

The estimation of freight trip generation is a critical component of traffic impact analyses (Holguín-Veras et al. 2013) and very important information to understand how goods move within an urban area. To Kulpa (2014, p. 197), freight trip generation refers to different areas or particular objects and may be estimated either by the number of vehicles (truck based model) or by the quantity of commodities measured in tons or values (commodity based model). Comi et al. (2012) also consider the delivery approach, which focuses on movements of goods measured in pick-ups and deliveries (delivery based model).

Brogan (1979) used trip rate per unit of area to understand the relationship between the number of truck-trips produced in/or attracted to an area, as well as the characteristics of that area. The author stated that the use of these rates contributes to the analysis of the impact of major truck-generating activities in certain sections of an urban area. An examination of the developed truck trip rates shows that, in general, commercial and industrial land uses are the largest generators of truck trips.

For Ogden (1992), understanding the relationship between transport and demand patterns is an essential component for modeling freight movement in urban areas. It evidences the link between movements of people and goods. Ogden (1992) also establishes a distinction between models based on movement of goods and models based on truck trips. In the model based on the movement of goods, producers and customers in a specific region create the demand for the movement of goods. Thus the movement of vehicles is a response to meet this demand. As a consequence, the truck trips establish a direct relationship with the movement of goods, and consequently the model outputs a certain amount of vehicle travel allocation. The author also developed models of attraction of products based on goods as well as on trips.

Tadi, Balbach (1994) estimated truck trips generated at warehouses, industrial sites, truck terminals and building areas using a regression model. Iding et al. (2002) developed a linear regression model that considers site area and number of employees as an alternative to the size of a business for each trip production and trip attraction per sector of industry.

Allen Jr. (2002) developed a demand forecast model to estimate the number of trips for medium and heavy trucks, considering the number of businesses and residences. Iding et al. (2002) present a freight trip model, obtaining attraction/production equations for different sectors, considering the area and the number of employees. Black (1999) developed similar study for Indiana-USA, considering 20 different products.

Comi et al. (2012) review the state-of-the-art on urban freight transport demand modelling and present different approaches after considering their pros and cons. One of the main conclusions of the authors is that the "direction relation between policies/measures and stakeholder's behaviours has not been investigated enough in the urban freight transport modelling literature" (Comi et al. 2012, p. 12). Anand et al. (2012) make a contribution to this gap, reviewing the literature concerned with urban freight modelling from the stakeholder's point of view describing their objective, descriptor choice and perspective towards the stakeholders' objective. Gonzalez-Feliu, Routhier (2012) analyze different models considering the function of the model, its scope and initial objectives, unit, model building approaches, choice of granularity and commercial applications and tools. An important conclusion of this review is about the relationship between goods flows and vehicle flows: if the focus is the impact of urban freight transport on urban traffic, the model unit must be considered vehicle-flows rather than commodity-flows which is commonly used in freight trip models. However converting commodity-flows to vehicle-flows is not a simple task (Gonzalez-Feliu, Routhier 2012). The authors claim that "modeling is not possible without knowledge of the existing situation, and in particular, knowledge about what is subject to change and what is invariable", so they suggest the development of procedures capable to understand the variation of data in space and time (Gonzalez-Feliu, Routhier 2012, p. 98).

Cherrett et al. (2012) indicate that 1.77 deliveries/week/100m^2 are transported in Winchester (UK). Using secondary data, Ducret, Gonzalez-Feliu (2016) attempt to connect a demand-estimation model and spatial modeling to evaluate the link between urban form characteristics and freight movements. The authors found that 2.9 deliveries/day are made by light commercial vehicles and 0.4 deliveries/day are made by single trucks, in Angers (France).

Holguín-Veras et al. (2016) discuss freight trip generation estimates. The authors indicate that "freight behavior research could be either a qualitative or a quantitative nature" (Holguín-Veras et al. 2016, p. 45). Priya et al. (2015) investigate an improved truck-trip model using time series, artificial neural networks and patterns of land-use.

With an innovative approach to model urban freight, Aditjandra et al. (2016) use a traffic survey and micro-simulation modeling to identify the impacts of freight deliveries and local service traffic at peak periods nearby the university campus in Newcastle (UK). The authors demonstrate that traffic surveys can be useful to assess the impacts of urban built environments on urban freight traffic. Sánchez-Díaz et al. (2016) propose a new set of variables that relate the establishment to its location, and assesses the

Table 1: Trip generation models to pubs and restaurants in Brazilian context

Reference	Equation	Trips per establishment	Dependent variable	R^2
Melo (2002)	T = 1.333 + 0.0019A	1.94	Area	0.797
Silva and Waisman (2007)	T = 2.963 + 0.0455A	-	Area	0.705
	T = 1.972 + 1.459E		Employers	0.739
	T = 2.074 + 0.017A + 0.967E		Area and Employers	0.754

Notes: T . . . Number of trips; A . . . Area of establishment; E . . . Number of employers

performance of novel explanatory variables on freight trip models. The results show that freight trip attraction is better modeled as a nonlinear function of employment and other geographic-based variables.

2.1 Urban freight models in the Brazilian context

Investigations of freight trip generation in the Brazilian context are incipient. There have so far been a few studies focused in São Paulo (Silva, Waisman 2007), Rio de Janeiro (Melo 2002, Gasparini 2008, Souza et al. 2010) and Campinas (Marra 1999). Marra (1999) developed a freight trip generation model designed for residential and commercial areas, based on the monthly freight flow. The model considers the area, the number of employees or residents and the average income per capita in the area as dependent variables. Melo (2002) developed an urban freight trip model for the city of Rio de Janeiro that considers groceries, clothing, retail stores, pubs and restaurants, construction material and fuel.

In particular, the sector of foods and beverages (focus of this paper) is a critical segment responsible for jobs, revenue and social activities vital to maintaining the economic sustainability of a city. Pubs and restaurants constitute a substantial volume of trips by attracting the movement of people and goods. Melo (2002) found 1.94 trips per day for establishments with an area between 60 and 780 square meters. Silva, Waisman (2007) developed a study that determined the freight trip generation rate for pubs and restaurants to a specific district in the city of Sao Paulo. The authors conducted a survey through a questionnaire for thirty establishments and collected information on daily freight trip rates, area of the establishment (in square meters) and number of employees. The authors identify the correlation between variables using multiple linear regressions. Gasparini (2008) developed a survey that analysed the freight trip rates for supermarkets and shopping malls in the city of Rio de Janeiro using the area, the number of customers, the number of parking spaces and the sales area as dependent variables in the model. The equations used in Brazilian freight trip studies to pubs and restaurants are presented in Table 1.

The approaches undertaken from the Brazilian context rely on the use of linear or polynomial regression and logarithm or exponential fit to determine the freight trip rates. Findings demonstrated a good fit ($R^2 \geq 0.7$), with wide divergences however, for the same variable as showed in Figure 1. Even with these good results, Souza et al. (2010) indicate the need for different equations or rates for different classes of goods and vehicles. This discussion demonstrates the complexity of the problem to be modelled and the importance of this topic of research.

The use of linear or polynomial regression as a methodology is a current trend, with Shin, Kawamura (2005), Bastida, Holguín-Veras (2009), Holguín-Veras et al. (2012), Holguín-Veras et al. (2013) and Priya et al. (2015) all using regression analysis to obtain a freight trip model. Sánchez-Díaz et al. (2016) incorporate locational variables and Aditjandra et al. (2016) use a traffic survey. The concept of locational variables is employed in the present study, which considers the importance of geographic-based information to evaluate which Brazilian model is more appropriate to estimate freight vehicle flows.

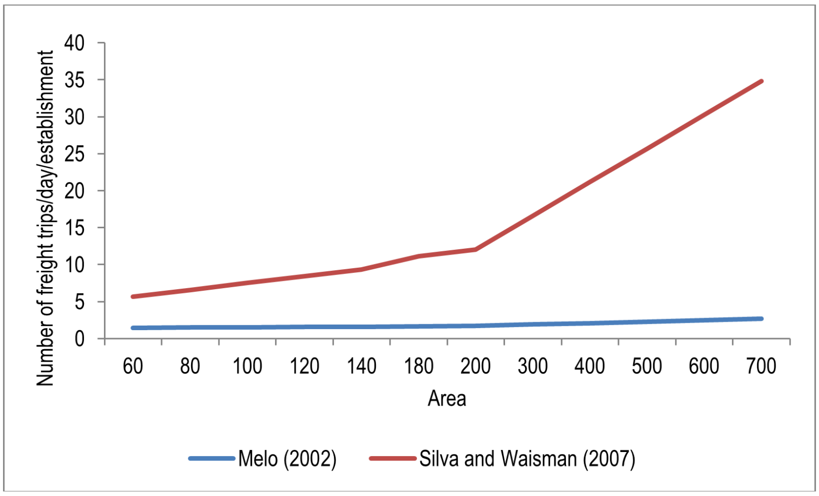


Figure 1: Graphical representation of Brazilian models to pubs and restaurants considering area as variable dependent

3 Study area and object of the investigation

This study focuses on the pubs and restaurants sector in Belo Horizonte city, which represents 42% of the number of deliveries in an area that contains the large majority of commercial establishments and urban mobility problems as well (Oliveira 2014, Oliveira, Guerra 2014). Belo Horizonte, founded in 1897 was planned to be the capital of the state of Minas Gerais. The plan envisioned a modern city that in 2010 could have a maximum population of 200,000 inhabitants living and working within the central boundary limited by Contorno Avenue (Oliveira et al. 2016).

Belo Horizonte has about 2.4 million inhabitants heterogeneously distributed over an area of 331 km^2. The city currently has the fourth largest Gross Domestic Product (GDP) of those in Brazil, with about 58.3 billion real (approximately currency rate: 1 euro is 3.60 real in October/2016). As a capital, commerce and public services are the key economic activities that propel the trading vocation of the city. The revenue comes mostly from services and commercial businesses, whereas the downtown area concentrates the federal, state and municipal governmental offices as well as an elevated number of commercial offices and industry headquarters. Belo Horizonte is surrounded by industrial-based cities and small towns mostly used as dormitory cities. The infrastructure of the capital's downtown makes it the most attractive area to retain hotels, restaurants, cafés and entertainment. This segment employs a substantial amount of jobs and makes an important contribution to the city's revenue.

In this investigation we hypothesize that a transportation model sensitive for geographic-based context will add quality and efficiency to computing freight trip generation by considering local characteristics. Thus, this paper analyses a freight trip generation model computed based on a survey developed for pubs and restaurants in the city of Belo Horizonte, Brazil. The distinctiveness of this study is the characterization of the sector and the geographical analysis of the results, which intends to inform transport policies for the distribution of food and beverage in pubs and restaurants. The methodology proposed in this study intends to provide a guideline for the planning of urban freight transport in order to obtain data for freight transport policies through field surveys (this type of analysis is still incipient in the Brazilian context).

4 Methodological approach

In theory, the data collection on movement of goods in an urban area can be obtained in different ways. In practice, however, because of the lack of waybill records or because of the nonexistence of registered information about trips, data about how, where the trips

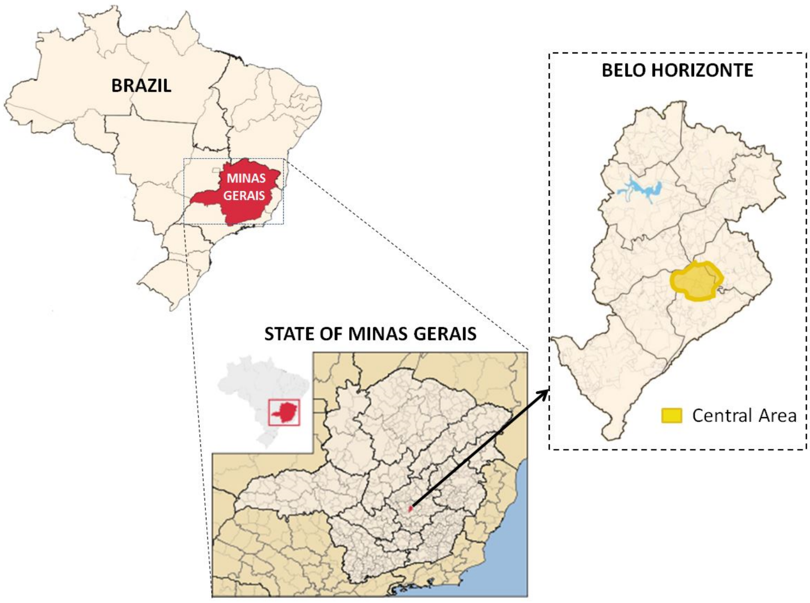


Figure 2: Geographic location of Belo Horizonte city and the study area

took place and about the volume of goods that flow in a city has been better obtained via survey. In this study we obtained data from specific stakeholders (retailers), based on interviews, as developed by Aditjandra et al. (2016). For the purposes of modeling urban goods, our data is categorized as vehicle/journeys (the route), as presented in Gonzalez-Feliu, Routhier (2012), and our model structure is defined as truck/vehicle as presented in Comi et al. (2012). These definitions are very important to characterize our methodology, which consists of:

1. develop a questionnaire to obtain freight flow data;
2. collect the data;
3. build and calibrate the model;
4. geographically analyse the results to understand the sector's impact on urban traffic.

4.1 Questionnaire

We developed a standard survey designed to obtain data that enables modeling of the vehicle journey categorized by foods and beverages. Table 2 presents the main information in the questionnaire. The questionnaire was validated with a pilot test in 10 commercial establishments and provides a database on the goods received in the establishments.

4.2 Data collection

The data collection was performed via interview. The interviews were conducted with the manager of the establishment or a technical staff member responsible for logistics operations. Unfortunately not all planned interviews were possible due to a lack of technical information, non-authorized personnel or when the establishment showed no interest in cooperating with the investigation.

The survey was conducted in 300 establishments. After the collection of data, the sample size was analyzed to determine the error and, for that, we used the simple random sample method. Other than information about the location of the pubs and restaurants, the interviewer also asked about desired and used routes, the time of the delivery and information about unload parking. Results are presented in Section 5.

Table 2: Main information of questionnaire

Type of information	Detailed information
Information of establishment	Location Name and function of interview Telephone and mail
Characteristics of establishment	Foundation time Day and working time Area (square meter) Number of employment
Receipt of goods	Origin of goods Number of trucks by day of week Time of delivery Unload time

4.3 Building and Calibrating the Model

The creation of the freight trip model presented in this paper was supported by a methodology developed by the Institute of Transportation Engineering (ITE 2008). Building a model requires a systematic comprehension of how data, tools methods and deliverables must be organized. Thus, the development of the model adds a methodological dimension for the approach, the data collection, the explanatory variables, the sampling elements and the date of the survey, that can be replicated in future investigations.

Besides the large number of variables and the complexity of the model for the development of the survey, it was also necessary to (ITE 2008):

- Consider the nature of the proposed model (type and size, geographical location, time/day that the model will be calibrated);
- Choose establishments that have characteristics that enables their travel demand to be modelled, such as:
  - it must be possible to reach an average rate of generation of trips with at least three elements;
  - it must be possible to use linear regression equations with six elements, but it is desirable to have at least 20 elements in the sample;
- Check the statistical correlation of the obtained model.

To define the model, we use linear regression. To validate the equations we use the coefficient of determination (R^2), which demonstrates how well an adjustment curve represents the relationship between the dependent and independent variables; the Student's t-test that determines the level of significance between the variables and the size of the sample; and the F-test that determines if the model is useful in predicting the flow. For this study, we considered valid the equations for which R^2 is greater than 0.5 (ITE 2008), ie, which represent 50% of the sample. In addition, we consider a confidence level of 95% for the t-test and significance level of 5% for the F-test. It is important to mention that models are best fit when R^2 is greater than 0.5, the sample size is greater than or equal to 4, and number of trips increases as the size of the independent variable increases (ITE 2008).

4.4 Geographical analysis

Once the trip generation model was computed, the establishments were located on the map by geocoding the addresses. We used Geographic Information System (GIS) tools to compute a continuous surface by interpolating the discrete values of travel per day delivered from the trip generation model. The method interpolates a raster surface from points using an inverse distance weighted (IDW) technique, which determines cell values using a linearly weighted combination of a set of sample points (travel per day per establishment). The weight is a function of inverse distance. The surface interpolated represents the dependent variable.

The interpolated surface represents the expected number of trips expanded to areas where sampling was not performed. Along with other maps, such as socioeconomics and transportation infrastructure, the spatial analysis helps to identify and understand the shortcoming and promotes a comprehensive standpoint for approaching the solution. For example, the solution can be applied to compute trip generation identifying areas with major demands of goods, therefore geographically correlating these areas with patterns of population income, age or behavior. This allows evaluation of the efficiency of the transportation policies that regulate urban distribution. As reported in Ducret, Gonzalez-Feliu (2016), which considered spatial categorization, urban freight modeling coupled to a GIS was used to demonstrate that spatial modeling can help organize logistics in the cities. In this work we aim to show the importance of integrating a robust freight trip generation model and GIS solution for providing an improved overview towards effective public policies for the city. Solutions in urban transportation planning must be addressed differently according to different aspects of population density and demand for goods.

5 Results and discussion

We used the Municipal Register of Taxpayers (MRT) provided by the Municipality of Belo Horizonte to know the total of establishments in the city. We identified 2,145 establishments as pubs, restaurants, juice houses or similar economic activity in this sector with the MRT database. We don't divide our sample and analyses in subcategories due to all establishments having similar characteristics: all commercialize food and drinks. Figure 3 shows the geographical dispersion of the selected establishments with high concentration in the 'Centro Sul' Region (59%). The 'Centro Sul' Region encapsulates the downtown area, and contains a large amount of jobs and a high density of residents, which compete daily for street use, parking spaces and mobility, therefore aggravating the problems of loading/unloading goods.

From the period July to August, 2014 we visited and interviewed three hundred establishments in the Belo Horizonte region. The sample has an error of 5.4% compared to the total amount of establishments, considering the simple random sample method. Our sample contains establishments with a building area of between forty-five and four hundred square meters, and with between one and thirty employees.

Our findings show that goods are delivered between five and six times a week in 66% of the establishments, except meat, fish and fresh food that have daily deliveries. The deliveries occur between 08:00 am and 10 am (71%). Beverages are delivered three to four times a week in 49% of the establishments, without a specific standard time for this kind of delivery. The peak day for deliveries is Tuesday (30%) followed by Friday (23%).

5.1 Freight trip generation model

After obtaining the data we applied the linear regression technique to correlate the dependent with the other independent variables. In this study, we considered as the dependent variable the number of freight trips and as the independent variable the area of the establishment and number of employees. We used the Pearson product-moment correlation coefficient to analyze the correlation between variables: when analyzing the number of freight trips and number of employees, , = 0.74 indicates a strong correlation between these variables. When considering number of freight trips and area of establishment; n = 0.70 also indicates a strong correlation between the variables. Using regression analysis to correlate the variables, we obtain results presented in Table 3. These models represent the number of daily freight trips (NT) attracted by establishments, considering area (A) and number of employees (E). It is necessary to multiply by a factor of 0.71 to determine the number of trips in peak hour (8am-10am).

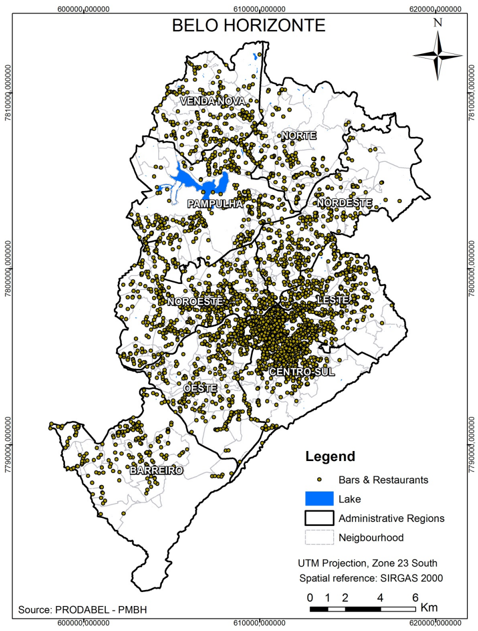


Figure 3: The geographical dispersion of Bars and Restaurants in Belo Horizonte

Table 3: Freight trip generation model to bars and restaurants

Model	Equation	R^2	T-Stat.	F-Test
Area (m^2)	NT = 0,98 + 0,0016A	0.48	1.97	283
	NT = 0,0076A	0.81	1.97	1,339
Employee	NT = 1,04 + 0,019E	0.55	1.97	363
	NT = 0,01E	0.67	1.97	605
Area (m^2) and Employee	NT = 1.01 + 0.014E + 0.00049A	0.55	1.97	189

The models obtained are valid for establishments with:

- Area between 40 and 400 m^2;
- Number of employees between 1 and 30.

The statistical significance of the results was tested: the squared correlation (R^2) indicates that a minimum of 48% of the variability in the "number of freight trips" is explained by the "area". The best model correlates the "number of freight trips" and "area".

Figure 4 shows the comparison of the results of this study with other Brazilian models. The comparison indicates that the Silva and Waisman model estimates a large number of freight trips and has lower accuracy. The results found in this study are similar to the model presented by Melo. To check the accuracy of the model, we compare the real data with the model outputs. Considering the model that correlates area and the numbers of employees (model 3), the results are 0.4% overestimated than the real values.

Figure 5 shows the concentration of deliveries in Belo Horizonte, considering freight traffic to deliver goods in pubs and restaurants between 8am-10am. These findings, when combined with the situation of the urban transportation infrastructure and the current transportation policies, reveals an alarming scenario. According to the numbers, in this scenario 71% of the pubs and restaurants require their goods deliveries in a very short time. Thus, it is easy to picture that this scenario creates a large amount of vehicles competing for parking spaces and reducing flow efficiency of the streets.

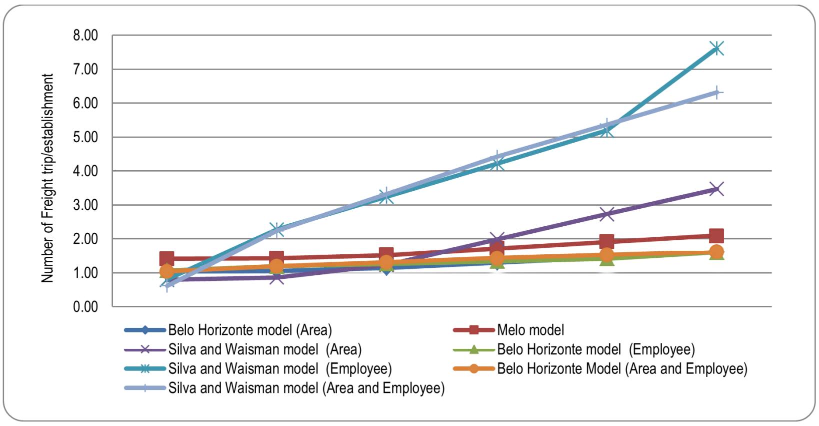


Figure 4: Comparison of models considering area (left) employers (center) and area and employees (right)

These results indicate the need for efficient public policies for urban freight transport. The pubs and restaurant sector is the only sector of the economy that generates a significant number of trips that overloads the transport infrastructure in these saturation times. Thus, the geographical analysis of results reinforces the results of the mathematical model and allows us to understand the extent of the problem.

5.2 Geographic overview of the model

The model provides an outstanding source for understanding the expected number of travels per business establishment. Based on this kind of information, urban planners and transportation managers can see and interact with indexes that reflect the demand for trip per business unit. However, the phenomena involved in the freight generation model includes a spatial component, which can be well explored if the geographic perspective is considered.

Transportation and economic development are areas intrinsically correlated in time and space. Therefore they rarely can be dissociated, especially when the goal of the study is to support public policies. For this reason, we introduced a basic structure, however realistic, for geographic analysis using the resulting model.

The model outputs the expected number of trips per day per establishment. The Municipal Register of Taxpayers data provides the addresses for the establishments. Together, these data were combined and displayed in a map. The surface provides the expected amount of trips per day for other regions in Belo Horizonte. It could be used, for example, to estimate trips per establishment for establishments that were not listed in the database or to provide an overview of the complex and delicate scenario of freight demand.

The geographic overview allows transportation and public managers to think about the freight in conjunction with other key factors that attract or repulse good deliveries in different regions of the city. Figure 6 illustrates a case scenario of downtown Belo Horizonte, which contains the majority of pubs and restaurants, a large amount of residents and a complex transportation network.

From the economic sustainability standpoint, there is an intrinsic relationship between population (consumers), foods and beverages (economic / social activity), mobility/accessibility to these services (existence of transportation infrastructure such as roads and public transportation) and the real demand for recharge supplies. The outputs of the model, when geographically spatialized and overlaid with other information provide an excellent perspective for supporting urban decision makers and policy makers.

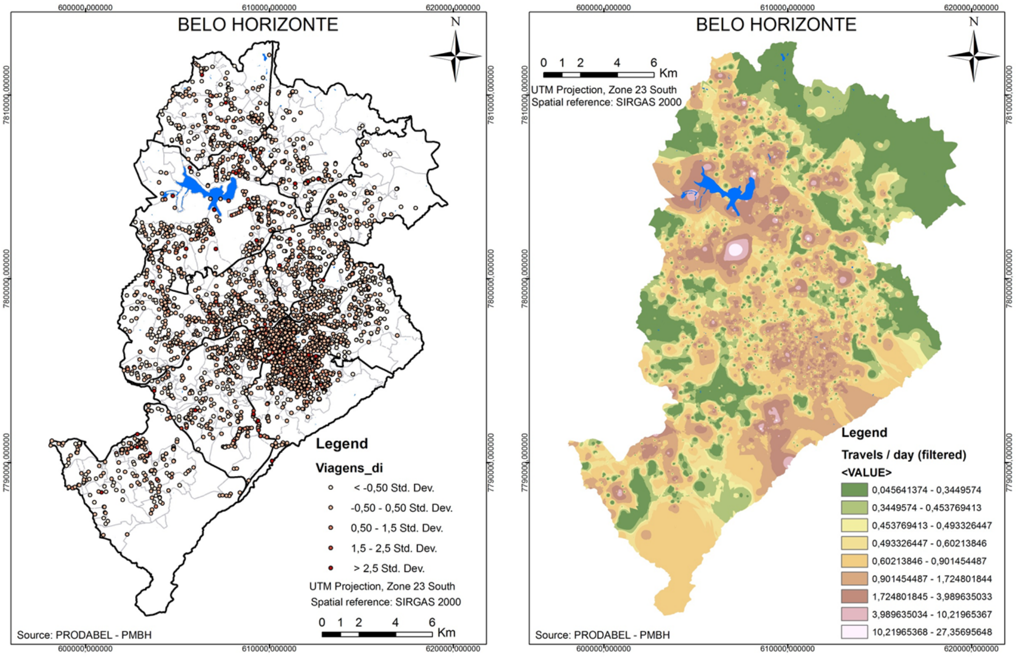


Figure 5: The geographical dispersion of deliveries to bars and restaurants in Belo Horizonte (left) and the surface of expected travel per day produced from the freight trip generation model (right)

5.3 Discussion of the results

The expected number of travels per day output from the model represents a realistic simulation of the current situation according to the survey from 2014. In order to better understand the results, the number of travels per day was loaded in a map containing the addresses of the establishments. This resulted in a punctual map (discrete information) that does not cover the whole area. Then, a geographic-based interpolation expanded the number of travels per day to a continuous surface as shown in Figures 5 and 6.

As shown in Figure 6, there is a geographic adherence between the number of travels per day, population and road density. The result indicates the need for an efficient urban freight policy to regulate the urban goods delivery towards meeting the demand of bars and restaurants.

As a recommendation, we draw attention to the importance of innovative ideas and alternative plans to lever positive changes in the urban transportation system to be presented and debated by specialists, authorities and the community. Holguín-Veras et al. (2016) suggest using forums and discussion groups to create a joint solution among the stakeholders involved in urban goods distribution. These measures highlight changes and can hopefully mobilize public managers, transportation managers and the population towards the continuous process of adaptation and improvement of public policies.

6 Conclusion

Regarding the development of this study, the key points are the characterization of an important economic sector of Belo Horizonte and especially the development of a freight trips generation model that is still an incipient theme in Brazil.

The model developed is mathematically robust and statistically acceptable. The model can be coupled to a GIS and provide results that support analysis in geographic contexts. We believe that the model can be used in the creation of public policies for Belo Horizonte: when we combine mathematical modeling with geographical analysis it is possible to understand the road occupancy impacts, and how and where the impacts will affect the population and economic sustainability of the region. These results can motivate the

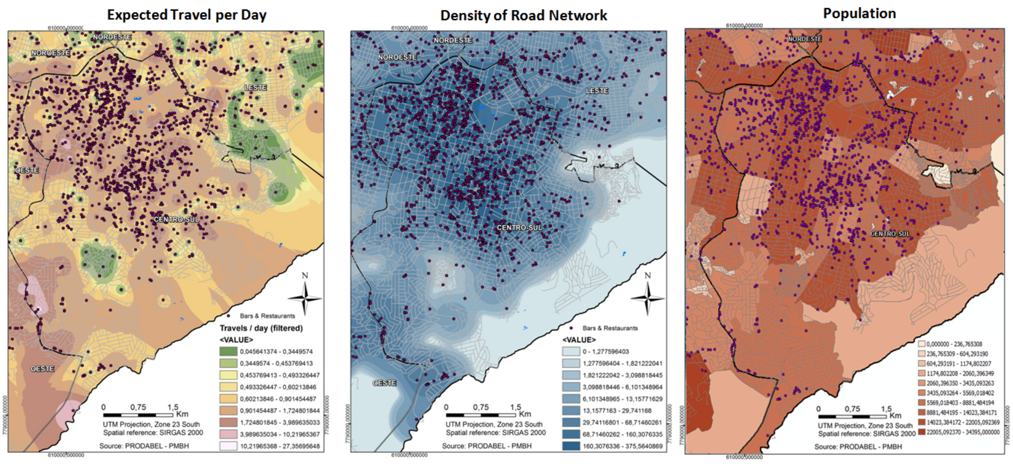


Figure 6: Overview of downtown Belo Horizonte using GIS maps used to improve the spatial analysis: the interpolated surface resulting from the trip generating model (left),the density of road network (center) and amount of population per neighborhood (right)

implementation of a less restrictive and more efficient public policy, to consolidate the goods and reduce the number of freight vehicles in urban areas.

The results of this study indicate differences between similar studies developed in Brazil. To do so, it demonstrates the necessity of the development of local models for freight demand management. Furthermore, the use of spatial analysis allows for the identification of potential sites for the implementation of this management policy, regarding the distances from main generators and attractors of freight trips.

References

Aditjandra PT, Galatioto F, Bell MC, Zunder TH (2016) Evaluating the impacts of urban freight traffic: Application of micro-simulation at a large establishment. *European Journal of Transport and Infrastructure Research* 16: 4–22

Allen Jr. WG (2002) Development of truck models. Report prepared for Baltimore Metropolitan Council, Baltimore, MD

Anand N, Quak H, van Duin R, Tavasszy L (2012) City logistics modeling efforts: Trends and gaps – A review. *Procedia - Social and Behavioral Sciences* 39: 101–115. CrossRef.

Bastida C, Holguín-Veras J (2009) Freight generation models: Comparative analysis of regression models and multiple classification analysis. *Transportation Research Record* 2097: 51–61. CrossRef.

Black WR (1999) Commodity flow modelling. Transportation research board. http://-onlinepubs.trb.org/onlinepubs/circulars/ec011/black.pdf

Brogan JD (1979) Development of truck trip-generation rates by generalized land use categories. *Transportation Research Record* 716: 38–43

Cherrett T, Allen J, McLeod F, Maynard S, Hickford A, Browne M (2012) Understanding urban freight activity – key issues for freight planning. *Journal of Transport Geography* 24: 22–32. CrossRef.

Comi A, Delle Site P, Filippi F, Nuzzolo A (2012) Urban freight transport demand modelling: A state of the art. *European Transport/Trasporti Europei* 51: 1–17

Ducret R, Gonzalez-Feliu J (2016) Connecting demand estimation and spatial category models for urban freight: First attempt and research implications. *Transportation Research Procedia* 12: 142–156. CrossRef.

Gasparini A (2008) Attractiveness of freight trip generators in urban areas. Master thesis. IME, Brazil. [in Portuguese]

Gonzalez-Feliu J, Routhier JL (2012) Modeling urban goods movement: How to be oriented with so many approaches? *Procedia - Social and Behavioral Sciences* 39: 89–100. CrossRef.

Holguín-Veras J, Jaller M, Sánchez-Díaz I, Wojtowicz J, Campbell S, Levinson H, Lawson C, Powers EL, Tavasszy L (2012) Freight trip generation and land use. National cooperative freight research program – NCFRP report 19. http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_739.pdf, accessed on 11 June 2015

Holguín-Veras J, Sánchez-Díaz I, Browne M (2016) Freight demand management: Role in sustainable urban freight system. *Transportation Research Procedia* 12: 40–52. CrossRef.

Holguín-Veras J, Sánchez-Díaz I, Lawson CT, Jaller M, Campbell S, Levinson HS, H-S S (2013) Transferability of freight trip generation models. *Transportation Research Record* 2397: 1–8. CrossRef.

IBGE – Instituto Brasileiro de Geografia e Estatística (2010) Population census 2010. Instituto brasileiro de geografia e estatística

Iding MHE, Meester WJ, Tavasszy LA (2002) Freight trip generation by firms. 42nd European congress of the European Regional Science Association, Dortmund, Germany

ITE – Institute of Transportation Engineers (2008) Trip generation handbook, an ITE recommended practice. Institute of transportation engineers, Washington, DC

Kulpa T (2014) Freight truck trip generation modelling at regional level. *Procedia - Social and Behavioral Sciences* 111: 197–202. CrossRef.

Marra C (1999) Chaaracterization of urban goods movements. Master thesis, UNICAMP, Brazil. [in Portuguese]

Melo IBC (2002) Demand assessment for freight transport in urban areas. Master thesis, IME, Rio de Janeiro

Ogden KW (1992) *Urban goods movement: A guide to policy and planning* (1st ed.). Ashgate Publishing, Burlington, VT

Oliveira LK (2014) Diagnosis of loading and unloading spaces for urban freight distribution: a case study in Belo Horizonte. *Journal of Transport Literature* 8: 178–209. CrossRef.

Oliveira LK, Guerra ED (2014) A diagnosis methodology for urban goods distribution: A case study in Belo Horizonte City (Brazil). *Procedia - Social and Behavioral Sciences* 125: 199–211. CrossRef.

Oliveira LK, Santos OR, Nóbrega RAA, Dablanc L (2016) The geography of warehousing in Belo Horizonte (Brazil). 14th World conference on transport research – WCTR 2016, Shanghai

Priya C, Ramadurai G, Devi G (2015) Freight trip generation models for Chennai, India. TRB 94th annual meeting compendium of papers, Washington, USA

Shin H, Kawamura K (2005) Data need for truck trip generation analysis: Qualitative analysis of the survey of retail stores. Transport Chicago, Illinois Institute of Technology

Silva MR, Waisman J (2007) Urban freight: an exploratory study about freight truck generation in bars and restaurants. 16th Brazilian congress of transport and traffic, Maceio, Brazil. [in Portuguese]

Sánchez-Díaz I, Holguín-Veras J, Wang X (2016) An exploratory analysis of spatial effects on freight trip attraction. *Transportation* 43: 177–196. CrossRef.

Souza CDR, Silva SD, D'Agosto MA (2010) Freight trip generators in urban areas. *Transportes* 18: 46–57. [in Portuguese]

Tadi RR, Balbach P (1994) Truck trip generation characteristics of nonresidential land uses. *ITE Journal* 64: 43–47

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 1, 2017, 31–49
DOI: 10.18335/region.v4i2.144
journal homepage: region.ersa.org

Creativity, Community, and Growth: A Social Geography of Urban Craft Beer

Neil Reid[1], **Jay D. Gatrell**[2]

[1] University of Toledo, Toledo, OH, USA (email: neil.reid@utoledo.edu)
[2] Bellarmine University, Louisville, KY, USA (email: jgatrell@bellarmine.edu)

Received: 15 July 2016/Accepted: 10 March 2017

Abstract. To better understand the non-economic drivers of growth in emerging industries, this paper examines the craft beer industry. Specifically, the paper will review two examples – the Black Cloister Brewing Company in Toledo, OH and Louisville's 3rd Turn Brewery – to understand how the values of entrepreneurs and local firms that are situated at the nexus of work, place, and creativity promote growth. Further, the paper will consider the socio-cultural meaning of creativity relative to the craft beer industry and the many ways in which the concept of innovation traditionally used by economic geographers to understand growth can be better understood within the context of creativity in select niche industries. In doing so, the paper represents a conceptual shift away from the linear process of innovation towards the more holistic notions of creativity, as well as community.

1 Introduction

> *Creativity is thinking up new things. Innovation is doing new things.*
> Theodore Levitt

The American craft beer industry has experienced significant growth in recent years. In November 2015, the number of craft breweries in the United States reached 4,144 eclipsing the previous high of 4,131 in 1873 (Brewers Association 2015d). This growth is particularly impressive considering that as recently as 2011 there were only 2,004 craft breweries in the United States (Brewers Association 2015b). Based on recent data, the growth is showing no signs of slowing down. Indeed, new breweries are opening at a rate of more than two per day and there are over 1,500 new breweries currently in the planning stages (Brewers Association 2015a). As a result, craft breweries now command a 12.2% share of the U.S beer market by volume and 21% by sales (Brewers Association 2015a, 2016a). Growth of the industry has been driven by consumer demand for alternatives to the beer that is produced by America's two mega-brewers – Anheuser-Busch and Miller Coors[1]. Yet, as we will explore, the growth of the craft beer industry – particularly at the local scale – may not simply be a desire for something different – or niche markets – but an industry intentionally embedded within a complex collection of non-economic

[1] On October 10 2016, a merger between Anheuser-Busch InBev and SABMiller was finalized. To comply with U.S. anti-trust regulations the combined company agreed to sell its stake in MillerCoors. MolsonCoors, who already had a 42% ownership stake in MillerCoors, purchased the remaining 58%. The purchase makes MolsonCoors the world's third largest brewing company (Nurin 2016, Svaldi 2016).

values and drivers – specifically creativity and a sense of community – that coincides with a passion for beer.

2 The US Beer Industry: From Macro to Micro

In order to appreciate the scale and scope of the disruptive force of craft beer on the broader U.S. brewing industry, it is essential to understand the historical context – specifically insofar as the sector has long been dominated by a few large macro-brewers with a narrow, or thin, collection of analogous beers. For example, the product line portfolio of Anheuser-Busch and Miller Coors is dominated by the American pale lagers (think Budweiser, Bud Light, Miller Lite, Coors) whose taste profiles are bland, undifferentiated, and unimaginative (Choi, Stack 2005). Anheuser Busch readily admits that over the years they have tinkered with the recipe of their signature Budweiser in order to make it less bitter and pungent, thereby making it more palatable to a larger and larger segment of the American population (Ellison 2006). The emergence of American beer as a bland homogeneous product is the result of the interplay of a number of events and trends that occurred in American society during the 20th century. These include Prohibition, the emergence and widespread adoption of refrigeration and packaging technologies, the utilization of television as a medium for mass marketing, the creation of national brands, consumer demand for homogenous products, brewery consolidation, and the emergent dominance of the mass production system that allowed brewers to realize economies of scale (Choi, Stack 2005, Tremblay, Tremblay 2009). For the large brewers, as their product became increasingly homogeneous, creativity was primarily restricted to the branding and marketing of their product lines (Imarenezor 2016, Monloss 2016).

While homogenous product lines produced by macro-brewers dominated the market throughout the 21st century, the seeds of disruption were sown in February 1978 with the legalization of home brewing and the subsequent state by state legalization of microbreweries and brewpubs. California and Washington were the first states to legalize brewpubs in 1982, with Oregon following the next year, 1983 (Fallows 2010). By 1995, brewpubs were legal in 45 states and the District of Columbia (Acitelli 2015). Today, they are legal in all 50 states. Home brewing is critical to the growth and success of the American craft beer industry. An estimated 1.2 million Americans brew beer at home. The two million barrels of beer that they produce on an annual basis represents approximately 1% of all the beer brewed in the United States. (American Homebrewers Association 2016). An estimated 90% of commercial craft brewer started out as home brewers. The websites of craft breweries are replete with stories of their founders starting off as home brewers (brewing beer in their basement, garage, or kitchen) and at some point making the decision to commercialize their hobby (Alonso 2011, Calagione 2011, Grossman 2013, Magee 2014). The home garage, kitchen, or basement is to the craft beer industry what the business incubator is for a high-tech start-up company.

The emergence and growth of the American craft beer industry is a clear indicator that increasing numbers of Americans are demanding more diversity in terms of their beer choices. They are no longer satisfied with drinking beers such as Budweiser or Miller Lite. The Brewers Association recognizes over 150 different styles of beer (Brewers Association 2015a), many of which are now produced commercially by American craft breweries. The result is that the American consumer now has an amazing variety of beers to choose from. This demand for craft beer is driven by a number of factors. One of these is neo-localism – the "deliberate seeking out of regional lore and local attachment by residents (new and old) as a delayed reaction to the destruction in modern America of traditional bonds to community and family" (Shortridge 1996, p. 10). As a result, there is a growing demand in the United States for products that have a strong connection with the community or region within which consumers reside. In concrete terms, neo-localism is manifested in the increasing numbers of farmers markets (Cone 2012), community supported agriculture initiatives (Galt et al. 2012), and community gardens (National Gardening Association 2014) across the United States. Craft beer, it has been argued, is yet another indicator of the strength of the American neo-localism movement. Schnell, Reese (2003, p. 66) suggest that "the explosive growth of microbreweries indicates a desire on the part of an increasing number of Americans, brewers and consumers alike, to reconnect with the cities or the towns in which they live, to resurrect a feeling of community tied to a specific landscape". Indeed there is evidence that "successful microbreweries are geographically focused, often producing specialized products with a strong local flavor" (Wesson, Nieva de Figueiredo 2001, p. 400). These are inherently local niche markets with limited range[2].

Based on market research, demand for craft beer appears to be driven primarily by the millennial cohort – young Americans between the ages of 21 and 34 (Crowell 2013, Duva 2014). Millennials have been characterized as "confident, liberal, upbeat, and open to change" (Pew Research Center 2010, p. 1). Millennials have higher levels of educational attainment than previous generations, prefer experimentation over brand loyalty, and are willing to pay more for a product that they perceive to be of higher quality (Pew Research Center 2010, Gilman 2014, Rotunno 2014). In making purchasing decisions an overwhelming majority of millennials (>80%) prefer to buy from companies that support solutions to specific social issues, while feeling that there is too much power concentrated in the hands of a few big companies (Winograd, Hais 2014). These values suggest a good match with craft brewers. Craft beer costs, on average, twice as much as macro-beer yet this does not seem to be squelching demand. Higher quality inputs are a significant contributor to the higher price point enjoyed by craft beer (Satran 2014). While large brewers engage in extensive philanthropy (e.g., see Anheuser-Busch 2016) craft breweries, because of their small size, are able to connect with their local communities in very customized and intimate ways. These include not only supporting local charities but also by making their brewery space available for community events and activities (Kirchenbauer 2014, The Beer Professor 2015). Craft breweries donate an average of $3.25 for every barrel of beer that they brew; the corresponding figure for Anheuser-Busch is only 35 cents per barrel (Herz 2014, Shilton 2016). Many craft breweries are also aggressively engaged in environmental initiatives designed to minimize their impact on the environment (McWilliams 2014, Schultz 2015, Ceres 2016). Environmental sustainability efforts include investing in renewable energy, waste recapture, recycling, and sustainable sourcing. As a result of these, and other, differences between craft and macro-brewers, we examine how the emergence of commercial craft brewing represents a departure from prior economic geographies and reflects a new social geography that is embedded with an emerging discourse on creativity, place, and values.

3 Theoretical Framework: Towards Creativity

Innovation has long been of interest to economic geographers. Indeed, geographies of high tech innovation, the spatial dynamics of innovation, and accounts of innovative places conceptually fueled much of the literature throughout the late-1980s and 1990s (Scott 1988, Storper, Walker 1989, Florida, Kenney 1990, Malecki 1991, Saxenian 1994, Lyons 1995, Audretsch, Feldman 1996, MacPherson 1997). Researchers focused on proximity and knowledge production, high tech corridors, and the diffusion of products and processes across space. This initial emphasis on innovation and technology gave rise to an explicit interest in science and technology inputs (such as engineering degrees) and the role of higher education in the development and deepening of the spatial division of labor.

Concomitantly, geographers sought to investigate sectors, including the service sector, which supported high-tech growth. These studies examined metropolitan regions and specific specialized industries (like so-called producer services) to understand the scale and scope of spatial interactions that define innovative communities, industries, and sectors (O hUallachain, Reid 1991, 1993, Markusen 1996). Eventually, geographies of innovation went micro and began to seek an understanding of the relationship between firm growth, knowledge production, and production systems on the ground and across space. In all of these studies, the research clearly demonstrated that the spatial division of labor and associated hierarchies, as well as "specialization", were well articulated, deepening, and that the cleavage between core and periphery was expanding.

[2] There are a small number of craft breweries (∼ 135) that have expanded beyond local markets and have a regional footprint of several dozen states (Reid, Gatrell 2015). These breweries represent ∼ 4% of the total brewery count.

Over time, geographies of innovation and the narratives of high tech growth came to dominate more applied economic development research, as well as the practice of economic development. As a result, communities across the U.S. and world sought to become the next "Silicon Valley", "Silicon Alley", or the "Silicon Forest". The emergence and expansion of science and technology driven growth was supported in large part by the popularity of Porter's Competitive Advantage of Nations (Porter 1990) and the so-called cluster approach. While Porter noted the existence of non-technological clusters – and Hayter (2000) has identified commodities-based clusters – the policy emphasis was almost exclusively on high tech growth and industries as communities struggled to revitalize local economies (i.e. Porter 1990, Scott 1993, Kaufman et al. 1994). Yet, the question of place always seemed to complicate innovation on the ground not only in peripheral regions with limited and often under-skilled labor markets, but also in larger metropolitan areas. In short, the 1990s and early 2000s illustrated that innovation couldn't be engineered in a global economy driven by the pernicious forces of uneven development per se – hence something else may have been driving the success of communities and differentiating places.

3.1 Beyond Linear Innovation Systems: Place & The Creative Class

Given the shortcomings of Porterian economic development, the spatial dynamics of neoclassical innovation models proffered by economists, and the explosion of research into the importance of the service sector, economic geography needed a new concept that reflected the intangible often-place based uncertainties that promote growth. Enter the abstract notion of "creativity" – innovation's non-linear socio-cultural and squishy partner and the creative class. Unlike innovation systems that could arguably be relocated or re-engineered, creativity as a concept and practice is inherently embedded in place and reflects a more holistic collection of socio-spatial relationships that are more cultural and historical – than a simple economic model. And these unique social-cultural relationships serve to define the creative class across space. The core of the creative class comprised elites, consumers, and leaders who shared progressive values and who drove urban revitalization in "hip" cities, established and emerging, across America.

According to Florida (2012), it is the creative class that drives and maintains the economic engine of cities vis-à-vis not only economic production; but the "squishy" concept of values. Like Supreme Court Justice Potter, Florida's new explanation for high growth regions boiled down to "I know it when I see it". The basic thesis asserted that urban growth and labor markets reinforce one another and highly articulated urban centers are anchored by a collection of elites that foster progressive values and an openness to new ideas and social change that attracts and retains talent. Florida's "creative class" was simultaneously a factor which explained production, consumption, and social change. Like any good economist, Florida and his colleagues asserted that the creative class could be modeled and as such quasi-quantitative indices such as the "bohemian index", "gay index", and "diversity index" proliferated the literature – and the popular media. Like the cluster before it, the "creative class" became a recipe that could be used to understand the comparative dynamics of economic development across space and gave rise to a full range of public policy initiatives – and associated progressive values.

Borrowing from Florida, we believe that emerging research on the "creative class" and progressive values has the potential to be an interesting and important conceptual development. That is, values can be used to explain economic development – and specifically a consumption driven model of development which deviates considerably from prior production centered frameworks like clusters. Moreover, the notion of the creative class is an attractive one, resonates with public discourses, and doesn't correspond to a single industry (i.e., specialization no longer dominates the discussion) and indeed economic diversification, entrepreneurism across multiple sectors, and niche industries were now as important as identifiable agglomerations such as "gun belt". In short, the new narrative could be used to revitalize the Rust Belt as well as explain the robust economies of tourist communities from Key West to Jackson Hole. Yet, the creative class doesn't explain creativity per se – it simply suggests you'll know it when you see it.

The purpose of this paper is to expand on the concept of creativity within a specific defined industrial niche. In contrast to work which emphasizes places, creativity as a concept can be examined by exploring a single industry – craft beer – that is closely associated with local markets, neighborhoods, and neo-localism (Flack 1997, Wesson, Nieva de Figueiredo 2001, Schnell, Reese 2003). Further, we assert that creativity is not synonymous with innovation or necessarily a process. Rather, creativity reflects a shared collection of socio-cultural relationships that might reside in place – or across a common industry, such as craft beer. For that reason, creativity is a highly contextualized set of values, interactions, and practices[3].

3.2 Space v. Place: Towards a Theory of Creativity & Community

As we have argued elsewhere, the everyday geographies of communities, culture, and politics of place are important to understanding the waxing and waning economics of localities (Gatrell, Reid 2002). To that end, we believe creativity – as a value – can be a useful framework which emphasizes niche local markets and entrepreneurism.

So what is the relationship between creativity and economic development from the perspective of geographers and regional scientists? Simply put – discussions of creativity and development extend from the Marxist concept of "creative destruction" as embodied in the classic work of Schumpeter (1942) (see also Scott 2006). In concrete terms, "creative destruction" asserts that the process of innovation in product and process is an ongoing regime that serves to eclipse the relevance of existing processes and products in the market place and can be used to explain capitalism's inherent spatial contradiction – uneven development (Smith 1984). While rooted in a Marxian reading of economics, "creative destruction" is not creativity per se. For that reason, we believe researchers have struggled to identify proxies, like patents, trademarks, industrial research, university research dollars and so on, to understand creativity in place.

More recently though, Törnqvist's The Geography of Creativity (Törnqvist 2011) provided researchers with a hybrid account of technological innovation that models the diffusion of technology across space and time, as well as the unique role 'special places' have played in fostering the innovation and creativity. Törnqvist's work echoes Scott (2006) insofar as creativity is a broader phenomenon embedded within social networks that are productive, consumption oriented, and symbolic. Yet, we would argue both Scott and Törnqvist are bound by the limited conceptualization of creativity as synonymous with technological innovation.

Unlike innovation, "creativity", as a spatial practice, gives meaning to place, production, and consumption – and the resulting "cultural" politics of local economic development (McCann 2002, Gatrell, Reid 2002, 2005). In the process, as Scott (2006) recognized – but framed differently - the practices associated with creativity serve to structure a "creative field" comprised of specific locational attributes and "geographically-differentiated webs of interaction" (Scott 2006, p. 3) that promote entrepreneurial activities – and enable entrepreneurs and creators to thrive. The classic example of so-called webs of interaction would be the growth, expansion, and sustained innovation observed in Silicon Valley. In the case of craft beer though, we are interested in understanding the values driven nature of the sector that resides at the nexus between creativity, community, and growth. Indeed, the local craft beer industry exists within a framework, or creative field, informed by shared experiences that drive economic development, transform neighborhoods, and may signal the potential to develop a new socio-cultural explanation for economic growth.

4 Craft Beer: Creativity as a Value

While the long-term viability of any business depends upon generating a profit there are some business owners for which non-financial returns enter into the calculus of their business model (Walker, Brown 2004). Such firms engage in what is termed satisficing behavior; in other words they are not driven by the maximization of profits. In contrast to a purely economic framework, satisficing behavior appeals to and obtains personal

[3] Hence, the notion that a "Creative Class" exists may within and between places be a misnomer. Rather, the discussions surrounding the creative class may reflect geography of progressive politics as it has become embedded within the U.S. urban hierarchy.

"value" from Maslow's hierarchy of needs (Maslow 1954). Specifically, Maslow (1954) proposes that other factors (such as physiological, safety, belonging, esteem, and self-actualization) influence behavior, motivate individuals, and inform decisions. In the business context, satisficing behavior occurs "when a firm's profit equals the managers' minimum acceptable or 'satisfactory' level and they then knowingly sacrifice additional profit in the pursuit of these higher order needs" (Kaufman 1990, p. 42). In the case of the craft beer industry, "belonging", "esteem", and/or achieving a firm's full creative potential (i.e., self-actualization) may take precedent over simple ROI. As these higher-order needs are non-monetary in nature they have been referred to as "psychic income" (Thurow 1978).

The idea that craft brewers engage in satisficing economic behavior has been hinted at by a number of researchers and writers. Wesson, Nieva de Figueiredo (2001, p. 392) observed that craft brewers are "often motivated by a love of brewing as by profits" while Day (2015) has described them as "astonishingly un-businesslike"[4]. A study of nascent craft brewers in Alabama concluded that "while the ultimate goal of brewers may be running a profitable operation, lifestyle or personal factors also appear to be important" Alonso (2011). Thurnell-Read (2014) examined the higher-order needs that seem to be valued by craft brewers in England. Among the brewers interviewed by Thurnell-Read there was "a sense of reward and satisfaction found in the production of what is perceived to be a product distinguished by the skills, passion and care deployed in its production." The brewers spoke of enjoying the creative process of making beer, the satisfaction they derived for being responsible for the final product, and the respect they felt from seeing customers drinking and clearly enjoying beer that they had brewed. These feelings were in sharp contrast to the boring, stressful, repetitive, and unfulfilling work that they had experienced in their previous jobs – descriptions that call to mind Marx's use of the adjectives alienation and estrangement to describe the worker's "relationship to the products of his labor" (Marx 1959, p. 30). In the documentary, Beer Wars (2010), Greg Koch, CEO of Stone Brewing Company, states that "I do not care how much beer we make. I only care how we make it". For the British, the selection of ingredients that go into their beer is driven by a desire to maximize product quality. Keeping down production costs is not part of the calculus.

Craft brewers also have a strong attachment to place. The ownership of craft breweries is ordinarily "local" – and entrepreneurial home brewers who decide to commercialize their hobby almost always do so in the place in which they live. The place often becomes a critical part in the identity of the brewery. The name of the brewery and the name of the beers that they produce often reflect local landmarks, historical figures, landscapes, historical events etc. (Flack 1997, Schnell, Reese 2003, Fletchall 2016). While most of the ingredients used by the breweries, with the exception of the water, are imported there is a growing interest in using more locally-grown ingredients. While three states (Idaho, Oregon, and Washington) grow 96% of the hops harvested in the U.S. there has, in recent years, been an expansion of hop growing outside of this core region. Driven by demand from craft brewers for local hops the plants are now grown in over 25 states (George 2016). While it is still the exception this has resulted in a small number of breweries producing beer made with all local ingredients. For example, in 2016 Boiler Brewing Company in Lincoln, Nebraska brewed a beer, appropriately named Nebraska Native, using all Nebraska ingredients (Matteson 2017). Examples such as this have been the catalyst for discussions around whether the concept of terroir has a place in the world of craft beer (Draft Magazine 2011, Bolden 2015). The connection that craft beer has with local places can be connected with the broader desire among growing numbers of people for food products that are locally-produced. As noted by Trubek, Bowen (2008, p. 24) "there is a growing movement that counters our placeless relationship to food and responds to increasing consumer demand for products that incorporate new dimensions of quality." Craft beer, made with high quality ingredients, by a brewer that lives locally,

[4]As craft beer has become a big business, global beer makers have sought to access craft and local markets through the strategic acquisition of micro-brewers, as well as larger regional craft brewers. Consequently, it could be argued that the current explosion of craft breweries – and the firm lifecycle – may be akin to the technology start-ups of the mid-1990s that enable small scale organizations to leverage modest novelty into sizable financial gains.

and with whom you can in all probability have a conversation with when you visit a local brewery is attractive to growing numbers of Americans.

Craft brewers are not only intimately connected to the product they make but also to the customers who drink their beer. These drinkers tend to be primarily millennials. A recent study by Mintel (2013) showed that 50% of older millennials (aged 25-34) drink craft beer compared with only 36% for the general population. Mike Stevens and Dave Engbers who founded Founders Brewing Company in Grand Rapids, Michigan note that "we don't brew beer for the masses. Instead, our beers are crafted for a chosen few, a small cadre of renegades and rebels who enjoy a beer that pushes the limits of what is commonly accepted as taste. In short, we make beer for people like us" (Founders Brewing Company 2016). The idea that craft brewers make beer for people like themselves should come as no surprise. People who make craft beer also drink craft beer.

The opportunity to be creative is valued by many craft brewers and is manifest in a number of different ways. Water, hops, barley, and yeast are the four basic ingredients that go into making beer. While the macro-brewers have a standard recipe upon which their beers are brewed craft brewers have creatively experimented with different types and combinations of these basic ingredients. Take hops for example. There are over one hundred different varieties of hops (Hopunion 2016). Different hops have different characteristics and the types and combinations used will influence both the flavor and aroma of the beer. In addition to different hop varieties craft breweries also use different hopping regimes that impact the flavor/aroma profile of their beer. Thus by adding hops at different stages of the brewing process (dry-hopping versus late-hopping) quite distinctive flavor/aroma profiles can be obtained (Schönberger, Kostelecky 2011). A number of craft brewers also engage in the practice known as wet-hopping which involves adding fresh hops (as opposed to say pelletized hops) that are delivered to the brewery within 48 hours of being picked (Vandenengel 2014). The plethora of hop varieties and the choice of different hopping regimes has been the catalyst for considerable creativity among craft brewers and have been credited with promoting new scientific research related to hops (Schönberger, Kostelecky 2011).

Beyond the four basic ingredients craft brewers are using their creativity to incorporate other ingredients into their beers. The list of these additional ingredients is almost endless and includes cacao, mango, cucumbers, tea and even bull testicles (All About Beer Magazine 2013, 2015a,b,d,e). Brewers are also working creatively to brew beers that no longer exist. Dogfish Head Brewery in Milton, Delaware have worked with Dr. Patrick McGovern an archaeologist at of the University of Pennsylvania Museum in Philadelphia to recreate a number of what they term ancient ales. These are ales whose recipes have been found in ancient tombs or derived from chemical analysis of ancient pottery (Dogfish Head 2016).

With so many craft breweries producing so many different styles of beer a major challenge facing many brewers is getting the consumer to notice and purchase their particular beers rather than all the others that are available. One way some craft brewers have addressed this challenge is to have creatively designed eye-catching packaging and labeling. A number of craft breweries hire brand consultants to assist them with the design of customized packaging and labels (American Craft Beer 2015). Brooklyn Brewery's iconic logo was designed by the celebrated New York graphic designer Milton Glaser (Agger 2013). In 2015 Arcadia Brewing Company of Kalamazoo, Michigan held a contest in which they invited local artists to submit label designs for their aptly named Art Hops brew (Mah 2016). On the other hand, some breweries opt for in-house design teams. In 2015 Ninkasi Brewing Co. of Eugene, Oregon announced the establishment of an artist-in-residence program whereby selected artists work with the brewery's in-house design and marketing team in the development of designs for new packaging and labeling (All About Beer Magazine 2015c).

5 Creativity, Passion, and Growth: Two Examples

As suggested above, economic geography's understanding of place, growth, and creativity have evolved over the past twenty years. As part of this evolution, the research suggests the scale of the analysis increasingly focuses on micro-level socio-economic contexts with special attention being paid to non-economic drivers of local growth. In an effort to understand the rescaling of local economic development, we believe the explosion of craft beer, as an industry, and the perspective of entrepreneurs active in this space are emblematic. For that reason, we examine two cases: The Black Cloister Brewing Company and 3rd Turn Brewing. The breweries were chosen because they are located in the cities in which the authors reside. This gave the authors easy access to the owners and staff for interview purposes and also allows them, through frequent visits as customers – a.k.a. participant-observation, to attain a feel for the atmosphere of the respective breweries. Likewise, the authors' insider knowledge as residents provide access to more local information resources and an understanding of the broader craft beer sector as it exists today and has developed over time in Toledo and Louisville. Additionally, both case studies are decidedly local insofar as the owners were homebrewers, intentionally invested in their home communities, and explicitly seek to create business that reflect a sense of place combined with business practices that are community centered.

5.1 The Black Cloister Brewing Company

The Black Cloister Brewing Company (BCBC) opened its doors in Toledo, Ohio in March 2015. The brewery is located in an 1874 structure that had been vacant since the 1970s in the heart of downtown Toledo (Cunningham 2015a). The brewery has the capacity to brew 1,000 barrels of beer per year which places it within the largest 1,000 craft breweries in the United States in terms of production volume. The CEO and Founder of Black Cloister is Tom Schaeffer. Schaeffer has an interesting background in that he is an ordained and practicing Lutheran Pastor. His Church, Threshold has its Sunday worship at the Black Cloister. Like many craft brewery owners Tom started out as a home brewer and was the founder of the local homebrew club, the Glass City Mashers (Brewers Association 2015c). Tom is also a certified Cicerone (the beer industry's equivalent of a wine sommelier).

Schaeffer (2016, personal interview, February 29) talked about the role that creativity plays at the Black Cloister, although he did so in quite different and distinctive ways. Schaffer began by talking about the role that creativity can play in helping old industrial cities like Toledo recover from the vagaries of the Great Recession of 2007-09. Toledo was particularly hard hit by the economic downturn and has since struggled to recover (Treanor, Costello 2015). In Schaeffer's opinion the retention and nurturing of creative workers is essential to Toledo's economic recovery and growth. Schaeffer notes that the Black Cloister has supported local creative businesses. The huge 50 foot mural that adorns the back wall of the brewery's tap room was designed by a local design and build studio, Graphite Design + Build, while the beer mugs that are given to members of the brewery's Mug Club are manufactured by a local glass blowing studio, Gathered Glassblowing studio. Both businesses are a few minutes' walk from the brewery. Schaeffer's mention of the importance of creative workers and industries resonates with the work of Richard Florida and his colleagues on the role of the creative class in economic development (Florida 2012, Lee et al. 2010). Additionally, in a highly informative paper on the economic history of Boston between 1630 and 2003, Glaeser (2006) emphasizes the importance of maintaining skilled workers during times of economic adversity. Boston was able to do that throughout its history whereas in cities like Detroit better educated, skilled, and creative workers abandoned the city. The most creative cities in the country, Schaefer noted, have a flourishing craft beer scene. This assertion is supported by Fallows (2016) who suggests that the existence of craft breweries in an indicator of a healthy and successful community. Other key characteristics of a healthy and successful community according to Fallows (2016) include a focus, by local residents, on working together, often in public-private partnerships to address local challenges, the existence of an identifiable civic champion, a compelling civic story, a vibrant downtown, the existence of a research university, a community college that is valued by and which serves the local community, elementary and secondary schools that have an element of experimentation about them, openness to outsiders, and a longer-term vision of what the community will look like in the future.

To Schaeffer creativity and brewing are linked. Brewing is very much a creative process. He makes a distinction between producing and creating beer. Producing beer is a process that Anheuser-Busch and Miller-Coors engage in. He recalled a conversation that he had with someone who had been a manager at a major brewery. The people there were hired to do a job, to produce beer, and not to engage in any creative thinking or production. Schaeffer also talks about the outcome of the creative process – the beer itself. There are times when he has made the decision to discard an entire batch of beer, not because there was anything wrong with it per se but because it did not match the intended flavor profile. The beer was perfectly fine (to use industry terminology there were no "off flavors") and could have been sold, at a significant profit, in the brewery's tap room. However, to have sold beer that did not match the intended flavor profile would have, in Schaeffer's opinion compromised the brewery's (and the brewer's) integrity. On some occasions, rather than discard beer, they have been able to sell it to a nearby distillery (Toledo Spirits) who have used it to produce, for example, hop-flavored whiskey. Even when it has been possible to sell a batch of beer to the distiller the brewery is making pennies on the dollar.

Head Brewer, Shannon Fink (2016, personal interview, February 9), had a rather serendipitous path into the world of professional brewing. She was a nurse for eleven years, a job that she disliked intensely. She found it unrewarding, stressful, emotionally draining, and overloaded with paperwork. In contrast, she describes her work environment at the brewery as "Zen-like". When asked to identify something that she disliked about working as a brewer she visibly struggled to identify anything and simply stated that she was "living the dream". Indeed, Fink enjoys the process of brewing so much ("It makes me happy" she stated) that on her days off she brews beer at home. Like Schaeffer, Fink started out as a home brewer. She had been introduced to home brewed beer by a nursing colleague and very shortly thereafter was brewing her own beer on her back patio. She was also educating herself by reading as many books about beer and brewing as she could lay her hands on. While her training and experience in nursing meant that she had a solid background in science, Fink also has an artistic streak. She taught herself how to draw in charcoal. She loved the one art course that she took in college. Her father, in fact, encouraged her to pursue her passion for art beyond college but Fink felt that there was no money to be made in art – until she entered the world of commercial brewing. She was a member of the local home brewing club that Schaeffer had founded and when the Black Cloister was close to opening he encouraged her to apply for one of the brewer's positions. Fink describes the brewing process as a "perfect marriage of art and science" (Cunningham 2015b). It is the creative process that she finds most rewarding – the opportunity to take ingredients and create a unique beer, a process of which she never tires. In recalling her days as a home brewer Fink described herself as being like a "mad scientist", always experimenting with different types and quantities of ingredients. Fink did not rely on reading books and trial-and-error to improve her brewing skills. She is also a graduate of the American Brewers Guild Craftbrewer's Apprenticeship Program, a 28-week program that focuses on brewing science and includes a five weeks of practical experience in the form of an internship at a craft brewery. It was after completion of this program that Schaeffer promoted Fink to the position of Head Brewer.

Fink also spoke about the satisfaction that she feels when she sees customers enjoying the beer that she has brewed. She often sits incognito at the Black Cloister's bar. Both Schaeffer and Fink talked about the role passion plays in the industry. Fink noted that much of the work that she does is not creative. She may, for example, spend an entire day cleaning various pieces of equipment. The job is also, at times, physically demanding as 50lb bags of malted barley have to be carried from one part of the brewery to another. But it her passion for the creative task of making beer that makes the less glamorous parts of the job worthwhile. Schaeffer spoke of the important part passion plays in the hiring of new employees. As part of the interview for a position at the Black Cloister interviewees are asked to talk passionately about a topic for five minutes. It can be any topic (it does not need to be beer). Schaeffer reckons that if an individual cannot talk passionately for five minutes on a topic about which they claim to be passionate then the probability is low that they will be able to talk passionately with customers about the brewery's beer. Knowledge of beer is less critical. As Schaeffer noted he can teach employees about beer; he cannot teach them to be passionate. As a result Schaeffer has listened to passionate monologues from interviewees on a variety of topics from dog parks to human genome sequencing.

5.2 3rd Turn Brewing

3rd Turn Brewing is located in Louisville, Kentucky. Like many craft beer firms, 3rd Turn's story is a familiar one that extends from an individual passion for beer and home brewing to a desire to create something meaningful – an experience and a product – in a local community (or place) that reflects the values and shared history of the owners. Opening in September 2015, the taproom initially focused 100% on guest taps. The micro-brewery served its first own commercial beer, a stout, in late-December. Going forward, 3rd Turn will continue to feature regional guest taps from more established Kentucky brewers such as Against the Grain, Apocalypse, Monnik, Beer Engine, and Country Boy. While many breweries limit outside pours, the owners are committed to celebrating the community of beer that has developed across the state and the Kentuckiana region. The tap room – is just that – a tap room. While no food is available for sale, the owners partnered with local restaurants for delivery, welcomes carry-ins (or potlucks), and the place serves as a gathering place.

3rd Turn was founded by three partners (Brian Minrath, Dale Shinkle, and Greg Hayden), each now on their third career. Despite an unintentional allusion to their personal narratives, the brewery was named for Churchill Downs' Third Turn and positions the brand squarely within the local history (Hayden et al. 2015, personal interview, December 22). Indeed, the name leverages local knowledge of Louisville insiders as the brand refers to a particular in-field location at Churchill Downs where residents can enjoy the derby which has now become an elite event for out-of-town guests and affluent locals. In this sense, the brewery, like many craft beer firms, builds on the well-established themes of neo-localism and explicitly appeals to a niche place-based market. Further, the vision of 3rd Turn blends the values of community insofar as the brewery owners sought "to create a space where J'Town [a community in metropolitan Louisville] residents and other East Enders" can enjoy beer "in a community setting." (Rothgerber 2016, p. 43).

What makes 3rd Turn's story interesting is that the endeavor represents an explicit and intentional desire on the part of Shinkle, Hayden, and Minrath, as well as Ben Shinkle (head brewer and Dale's brother), to re-vision their lives as entrepreneurs and to create something meaningful for their families and hometown. Lifestyle, values, and meaning-making drive their entrepreneurism. For example, Dale and Greg are both chemists who've been friends since college and both subsequently pursued careers in business as a CPA and safety planner, respectively (Hayden et al. 2015, personal interview, December 22). It was in the corporate world that Greg met their third partner, Minrath. As of early 2016, Greg was the only founder that had transitioned to full time at the brewery to pursue his passion and new career. In the case of Ben, Greg, and Dale, the 3rd Turn venture permitted them to take their hobby to the next level. While Ben is the head brewer, 3rd Turn specializes in the scaled up recipes developed by the former chemists – in particular Greg – that serve as the foundation of the production system. To that end, Greg and Dale's passion for chemistry and beer continues to inform their work – and each speaks of esters, sugars, and active yeast with a passion that reflect their desire to create "something new".

In terms of their beer, 3rd Turn explores and creates beers that intentionally reside outside of the traditional style guide. For example, the team created a hefe gerst based on a Ukrainian Steam Beer that uses barley as the primary grain (Shinkle 2016, personal interview, February 27). With a current annual maximum production capacity of roughly 2,000 barrels, the goal is to create unique beers in much smaller quantities that will be served primarily at the tap room – as well as on guest taps across Kentucky and southern Indiana[5]. The theme of creativity and collaborating with the community also

[5] In late December 2016, 3rd Turn announced a planned expansion and new facility in nearby Crestwood, KY located in Oldham County. The new "farm" location will permit the firm to expand production.

extends to their "mug club". Like many tap rooms, 3rd Turn has implemented an annual membership program and the mugs are handcrafted by a local artist in residence, Fong Choo – a faculty member at nearby Bellarmine University which is also Dale and Greg's alma mater. The individual nature of the ceramic mugs makes 3rd Turn unique and reflects the owner's desire to seek inspiration, foster community, and develop collaboration in non-traditional arenas. Like BCBC, 3rd Turn is operationalizing creativity broadly and deploying everyday strategies that reinforce economic development and themes from the "creative class".

Beyond creating beer, the owners clearly value community, hard work, collaboration, and fun. However, they understand business too. Indeed, site selection was a lengthy process that included working with local officials in multiple locations in suburban Louisville to identify an underserved market (think intervening location). As a result of their efforts, they chose a location – a former Moose Lodge that was once a church – which enabled 3rd Turn to be the first brewery located beyond the outer belt – miles away from other more stablished local competitors – many of whom have pours available on 3rd Turn taps.

6 Discussion

Craft brewers and beer drinkers are part of a broader local movement that increasingly redefines, or perhaps reframes, local economic development. Or put another way, the values-oriented entrepreneurs driving the craft brew industry see themselves as engines of community-building at the micro-scale. In doing so, craft brewers become "change agents". As change agents, brewers and breweries are often the anchor of revitalizing neighborhoods and spur further capital investment. In both of the case studies, the firms intentionally located in more marginal or transitional areas and the breweries are adaptations of prior uses. Moreover, the realities of the craft brewers – and the micro-spaces which they occupy are instructive as their success depends heavily on their ability to leverage creativity (product, marketing, and brand identity) in a very dense marketplace (Mann 2016).

According to Julia Herz, craft beer program director for the Brewers Association, craft "brewers are not just opening up businesses to run a profit; they have a different version of the American Dream – one where you can use your brewing as a platform for improving your community." (Shilton 2016). Craft breweries both contribute to and benefit from agglomeration economies. In an increasing number of cases these agglomeration economies are buttressed by geographic concentrations of locally-owned, often, creative businesses. Indeed, breweries are sometimes seen by policy makers as one of the critical elements of economic growth and neighborhood revitalization (City of Louisville 2014, Loosemore 2016). Indeed, the location of the vast majority of craft breweries in Louisville coincides with revitalizing (arguably artistic or bohemian) urban districts such as Portland (Against the Grain production, Falls City), NuLu (Against the Grain, Akasha and Goodwood), and Germantown (Monnik). To that end, the city of Louisville has specifically targeted the craft beer industry as a critical element of growth in select neighborhoods and a working group has proposed a bike-able "beer trail" similar to the successful bourbon trail (City of Louisville 2014). In the case of Toledo's Black Cloister Tom Schaeffer noted the geographically proximate glass blowing and build and design studios that have contributed to the ambience that he is trying to create for his customers.

In another corner of Ohio, the Great Lakes Brewing Company has been credited with being a catalyst in the socio-economic revitalization of the Cleveland's Ohio City neighborhood. In the late nineteenth and early twentieth century Ohio City was a bustling industrial neighborhood with docks, mills, foundries, distilleries, and bottling works (Ohio City Neighborhood 2016c). By the end of the Second World War, however, the neighborhood was in decline. Pat and Dan Conway, who founded the Great Lakes Brewery in 1986, have been described as the "beer men who became the unlikely leaders of the neighborhood's revival" (Alexander 2013). Today, Ohio City is a vibrant, ethnically diverse neighborhood (34% of its 9,000 residents are African American and 23% are Hispanic) that is teeming with small locally-owned businesses. These include other breweries, restaurants, bicycle shops, book stores, a glass blowing studio etc. (Ohio City

Neighborhood 2016b,d). Since 2005, the crime rate has fallen approximately 25% and real estate values have doubled. It has been rated as the second most walkable neighborhood in Cleveland (Ohio City Neighborhood 2016a).

Like Ohio City, Denver's Lower Downtown (LoDo) neighborhood is another that has benefited from the establishment of a craft brewery. In this case it is the Wynkoop Brewery. Like Ohio City, Lower Downtown Denver was once a bustling neighborhood and a major industrial area that drew its sustenance from the Union Pacific rail yard. By the early twentieth century LoDo was in decline as the railroad's importance as transportation medium decreased. It was not until the late-1980s that LoDo started to witness a revival and the Wynkoop Brewery played a key part in that process. "LoDo's characteristics make it an ideal site for the downtown's restaurant and entertainment center. The Wynkoop was the pioneer which laid the foundation for this vision's realization, which then became an anchor for the broader development of LoDo" (Weilar 2000, p. 175). Today, much like Ohio City, the area is thriving and is home to some of the best restaurants, galleries, shops and boutiques in the city, as well as a dozen or so craft breweries. Compared with Ohio City, however, the LoDo neighborhood is not as demographically diverse; 80.4% of its population is white. Out of 32 Denver neighborhoods it is the 27th in terms of the share of non-whites in its population (Statistical Atlas 2015).

As the discussion suggests, local craft brewers are critical "place-makers" and contribute to the overall sense of place associated with Millennial driven urban revitalization.[6] This is a process and a relationship that is well-established and shows little sign of slowing down. At the same time this has been very much an organic process with home brewers making the choice to commercialize their hobby and in so doing making an incremental contribution to local economic development. However, as we have seen in the Louisville case, the growth of craft brewing has caught the attention of policy makers with the result that some communities are starting to think strategically about how to foster and nurture the industry. This includes enacting legislation to make the regulatory environment friendlier to the industry (Bragg 2016).

7 Conclusion

While it may be too much to say that the craft beer industry has disrupted the broader beer industry there is no doubt that its emergence has impacted in a significant way the large multi-national brewing companies. Faced with increased competition from craft brewers in the United States, Anheuser-Busch and SABMiller finalized a $100 billion merger in 2016. This merger provides unprecedented economies of scale and, at a time when its share of the U.S. market is declining, gives Anheuser-Busch access to growing markets in Africa and Latin America (Brown 2016). Craft beer's emergence in the United States may now be actively re-scaling economic development and reconfiguring traditional explanations of economic growth. While growth may be a by-product of craft beer entrepreneurs, the evidence suggests the drivers and decisions of entrepreneurs are nuanced, more local, and often reside outside of the simple arithmetic of ROI. While no doubt all entrepreneurs seek to maximize their investment, non-economic drivers such as creativity, community, and even a passion for brewing are critical to the success of craft firms. Indeed, the success of the firms themselves is largely dependent upon the creativity of brewers and the neighborhoods (or communities) within which they are embedded.

While the notion of values-driven entrepreneurism may not necessarily be new (e.g., Ben & Jerry's or Hobby Lobby), the experience of the craft beer industry is unique as the framework that drives the sector may be signposting the emergence of new hybrid models of economic development that account for place, community, creativity, and growth. As the two cases illustrate, the industry at the local scale appears to have been driven – in part – by non-economic drivers and that these drivers are inherently scaled, as well as experienced, at the everyday world of neighborhood communities. More importantly, and

[6] While craft brewing was popular before the appearance of the so-called millennial cohort research and data suggest that the millennials are the key demographic that have driven growth of the industry more recently. Recent market research shows that 57% of people who drink craft beer on a weekly basis are millennials, compared with 24% and 17% for Gen Xers and Baby Boomers, respectively (Brewers Association 2016b).

not surprisingly, the language of craft brewers parallels that of the creative class and reflects the shifting consumption patterns and lifestyle of a now established millennial cohort.

References

Acitelli T (2015) When brewpubs started booming. *All About Beer Magazine*, February 11. http://allaboutbeer.com/when-brewpubs-started-booming/

Agger M (2013) I love Brooklyn Brewery. *The New Yorker*, March 5. http://www.-newyorker.com/culture/culture-desk/i-brooklyn-brewery

Alexander D (2013) Beer entrepreneurs fuel comeback of struggling Cleveland neighborhood. *Forbes*, November 26. http://www.forbes.com/sites/danalexander/-2013/11/26/beer-entrepreneurs-fuel-comeback-of-struggling-cleveland-neighborhood/-#14cc94c3574f

All About Beer Magazine (2013) Join all about beer for a beer tasting with Wynkoop Brewing Co. *All About Beer Magazine*, March 29. http://allaboutbeer.com/news/beer-tasting-with-wynkoop-brewing-co/

All About Beer Magazine (2015a) Cucumber beers: stars of the summer? *All About Beer Magazine*, August 13. http://allaboutbeer.com/cucumber-beers/

All About Beer Magazine (2015b) Founders Brewing Co. announces first backstage series beer of 2016: 'mango magnifico'. *All About Beer Magazine*, January 25. http://allaboutbeer.com/news/founders-brewing-backstage-series-2016-mango-magnifico/

All About Beer Magazine (2015c) Ninkasi Brewing Co. announces artist in residence program. *All About Beer Magazine*, June 26. http://allaboutbeer.com/news/ninkasi-brewing-co-announces-artist-in-residence-program/

All About Beer Magazine (2015d) Stone brewing re-releases stone 12th anniversary bitter chocolate oatmeal stout. *All About Beer Magazine*, January 25. http://allaboutbeer.com/news/stone-12th-anniversary-bitter-chocolate-oatmeal-stout/

All About Beer Magazine (2015e) Tea: an untapped beer ingredient? *All About Beer Magazine*, April 23. http://allaboutbeer.com/tea-an-untapped-beer-ingredient/

Alonso AD (2011) Opportunities and challenges in the development of micro-brewing and beer tourism: A preliminary study from Alabama. *Tourism Planning & Development* 8[4]: 415–431

American Craft Beer (2015) The art of branding craft beer. *American Craft Beer*, February 18. http://americancraftbeer.com/item/the-art-of-branding-craft-beer.html

American Homebrewers Association (2016) Homebrewing stats. http://www.homebrewers-association.org/membership/homebrewing-stats/

Anheuser-Busch (2016) Serving our communities with pride. http://www.anheuser-busch.-com/betterworld/community.html

Audretsch D, Feldman M (1996) R&D spillovers and the geography of innovation and production. *The American Economic Review* 86[3]: 630–640

Beer Wars (2010) Dir. Anat Baron. Ducks in a Row Entertainment Corporation. DVD

Bolden E (2015) Can craft beer truly express a sense of place? *Punch*, July 9. http://punchdrink.com/articles/can-craft-beer-express-terroir-with-rogue-farms-wolves-and-people-brewery/

Bragg V (2016) New bill proposed to help attract more craft breweries to Virginia. *WSET.com*, January 8. http://wset.com/news/local/new-bill-proposed-to-help-attract-more-craft-breweries-to-virginia

Brewers Association (2015a) Brewers association reports big gains for small and independent brewers. July 27. https://www.brewersassociation.org/press-releases/brewers-association-reports-big-gains-for-small-and-independent-brewers-2/

Brewers Association (2015b) Number of breweries. https://www.brewersassociation.org/-statistics/number-of-breweries/

Brewers Association (2015c) Tom Schaeffer. https://www.brewersassociation.org/profile/-tom-schaeffer/

Brewers Association (2015d) The year in beer: U.S. brewery count reaches all-time high of 4,144. December 2. https://www.brewersassociation.org/press-releases/the-year-in-beer-u-s-brewery-count-reaches-all-time-high-of-4144/

Brewers Association (2016a) Small and independent brewers continue to grow double digits. March 22. https://www.brewersassociation.org/press-releases/small-independent-brewers-continue-grow-double-digits/

Brewers Association (2016b) Today's craft beer lovers: Millennials, women and hispanics. August 15. https://www.brewersassociation.org/communicating-craft/understanding-todays-craft-beer-lovers-millennials-women-hispanics/

Brown L (2016) A-B InBev finalizes $100b billion acquisition of SABMiller, creating world's largest beer company. *Chicago Tribune*, October 11. http://www.chicagotribune.com/business/ct-megabrew-ab-inbev-sabmiller-merger-20161010-story.html

Calagione S (2011) *Brewing Up A Business: Adventures In Beer*. John Wiley & Sons, Hoboken, NJ

Ceres (2016) Brewery climate declaration. http://www.ceres.org/declaration/about/climate-declaration-campaigns/brewery#breweries

Choi DY, Stack MH (2005) The all-american beer: A case of inferior standard (taste) prevailing. *Business Horizons* 48[1]: 79–86. CrossRef.

City of Louisville (2014) Mayor's local brewery workgroup report. City of Louisville, Louisville, KY. http://louisvilleky.gov/sites/default/files/mayors_office/pdf_files/-mayors20beer20report20201420ver202.pdf

Cone T (2012) Number of US famers markets surges. *USA Today*, August 3. http://usatoday30.usatoday.com/news/nation/story/2012-08-03/farmers-markets-increasing/56727330/1

Crowell C (2013) Craft beer consumer stats: How will they affect your business plan? *Craft Brewing Business*, May 3. http://www.craftbrewingbusiness.com/business-marketing/craft-beer-consumer-stats-how-will-affect-your-business-plan/

Cunningham B (2015a) Black Cloister Brewing Co. taps opening date. *Toledo Blade*, February 26. http://www.toledoblade.com/Peach-Weekender/2015/02/26/Black-Cloister-Brewing-Co-taps-opening-date.html

Cunningham B (2015b) Black Cloister's Fink is head of the class. *Toledo Blade*, September 24. http://www.toledoblade.com/Peach-Weekender/2015/09/24/Black-Cloister-s-Fink-is-head-of-the-class.html

Day P (2015) How the craft beer revolution started. *BBC News*, December 30. http://www.bbc.com/news/business-35120401

Dogfish Head (2016) Ancient ales. http://shop.dogfish.com/ancientales

Draft Magazine (2011) The dirt on terroir. *Draft Magazine*, July 26. http://draftmag.-com/the-dirt-on-terroir/

Duva N (2014) Elitism, or something else? Millennials and the war on big beer. *CNBC.com*, November 8. http://www.cnbc.com/2014/11/07/why-those-elitist-millennials-hate-big-beer.html

Ellison S (2006) After making beer ever lighter, Anheuser faces a new palate. *Wall Street Journal*, April 26. http://www.wsj.com/articles/SB114601602889736048

Fallows J (2010) Jimmy Carter: Not the king of beers? (updated). *The Atlantic*, August 19. http://www.theatlantic.com/politics/archive/2010/08/jimmy-carter-not-the-king-of-beers-updated/61599/

Fallows J (2016) Eleven signs a city will succeed. *The Atlantic*, March. http://www.-theatlantic.com/magazine/archive/2016/03/eleven-signs-a-city-will-succeed/426885/-?.utm_source=SFFB

Fink S (2016) Personal interview. Toledo, OH, February 9

Flack W (1997) American microbreweries and neolocalism: "Ale-ing" for a sense of place. *Journal of Cultural Geography* 16[2]: 37–53. CrossRef.

Fletchall AM (2016) Place-making through beer-drinking: A case study of Montana's craft breweries. *Geographical Review* 106[4]: 539–566. CrossRef.

Florida R (2012) *The Rise of the Creative Class, Revisited.* Basic Books, New York

Florida R, Kenney M (1990) Silicon Valley and Route 128 won't save us. *California Management Review* 33[1]: 68–88. CrossRef.

Founders Brewing Company (2016) History. http://foundersbrewing.com/history/

Galt RE, O'Sullivan L, Beckett J, Hiner CC (2012) Community supported agriculture is thriving in the central valley. *California Agriculture* 66: 8–1. CrossRef.

Gatrell J, Reid N (2002) The cultural politics of local economic development: The case of Toledo Jeep. *Tijdschrift voor Economische en Sociale Geografie* 93[4]: 397–411. CrossRef.

Gatrell J, Reid N (2005) Placing economic development narratives into emerging economic spaces: Project Jeep, 1996-1997. In: LeHeron R, Harrington J (eds), *New economic spaces: New economic geographies.* Ashgate, Aldershot

George A (2016) 2016 world harvest report U.S. hops. *The New Brewer* November/December: 64–69

Gilman HR (2014) How millennials' use of social networks explains their politics. *Brookings Tech Tank*, November 24. https://www.brookings.edu/blog/techtank/2014/11/24/how-millennials-use-of-social-networks-explains-their-politics/

Glaeser EL (2006) Reinventing Boston: 1630-2003. *Journal of Economic Geography* 5: 119–153. CrossRef.

Grossman K (2013) *Beyond the pale: The story of the Sierra Nevada Brewing Co.* John Wiley & Sons, Hoboken, NJ

Hayden G, Shinkle D, Shinkle B (2015) Personal interview. Jeffersontown, KY, December 22.

Hayter R (2000) *Flexible Crossroads: The Restructuring of British Columbia's Forest Economy.* UBC Press, Vancouver, BC

Herz J (2014) Craft brewers' 2014 charitable donations top $71m. *Brewers Association* https://www.brewersassociation.org/communicating-craft/craft-brewers-contribute-millions-to-local-causes/

Hopunion (2016) A hop for every flavor and style. https://www.hopunion.com/hop-varieties/

Imarenezor C (2016) Super Bowl 50: Anheuser-Busch returns to commercial lineup for 28th year. *Vibe*, January 16. http://www.vibe.com/2016/01/super-bowl-50-anheuser-busch/

Kaufman A, Gittell R, Merenda M, Naumes W, Wood C (1994) Porter's model for geographic competitive advantage: The case of New Hampshire. *Economic Development Quarterly* 8[1]: 43–66. CrossRef.

Kaufman BE (1990) A new theory of satisficing. *The Journal of Behavioral Economics* 19[1]: 35–51. CrossRef.

Kirchenbauer H (2014) Philanthropy is alive and well in the craft beer industry. *San Diego Troubadour*, June. http://sandiegotroubadour.com/2014/06/philanthropy-is-alive-and-well-in-the-craft-beer-industry/

Lee SY, Florida R, Gates G (2010) Innovation, human capital, and creativity. *International Review of Public Administration* 14[3]: 13–24. CrossRef.

Loosemore B (2016) Falls City announces new bottled beers. *The Courier Journal*, March 1. http://www.courier-journal.com/story/life/food/spirits/beer/2016/03/01/falls-city-announces-new-bottled-beers/81158592/

Lyons D (1995) Agglomeration economies among high technology firms in advanced production areas: The case of Denver/Boulder. *Regional Studies* 29[3]: 265–278. CrossRef.

MacPherson A (1997) The role of producer service outsourcing in the innovation performance of New York state manufacturing firms. *Annals of the Association of American Geographers* 87[1]: 52–71. CrossRef.

Magee T (2014) *So you want to start a brewery? The Lagunitas story.* Chicago Review Press, Chicago, IL

Mah L (2016) Art Hops beer and logo design make debut at Friday night Art Hop. *Mlive*, January 9. http://www.mlive.com/entertainment/kalamazoo/index.ssf/2016/01/art_hops_beer_and_logo_design.html

Malecki E (1991) *Technology and Economic Development: The Dynamics of Local, Regional, and National Change.* Longman, New York, NY

Mann D (2016) Five funny names in Louisville beer – and why brewers come up with them. *Louisville Business First*, February 1. http://www.bizjournals.com/louisville/news/2016/02/01/five-funnynames-in-louisville-beer-and-why-brewers.html

Markusen A (1996) Sticky places in slippery space: A typology of industrial districts. *Economic Geography* 72[3]: 293–313. CrossRef.

Marx K (1959) *Economic and philosophical manuscripts of 1844.* Progress Publishers, Moscow

Maslow AH (1954) *Motivation and Personality.* Harper & Row, New York, NY. CrossRef.

Matteson C (2017) Boiler brewing company tackles Nebraska-made beer. *Lincoln Journal Star*, January 11. http://journalstar.com/lifestyles/food-and-cooking/boiler-brewing-company-tackles-nebraska-made-beer/article_bfe16c27-b68d-50d2-a90a-fa13fda7ff89.html

McCann E (2002) The cultural politics of local economic development: Meaning making, place-making, and the urban policy process. *Geoforum* 33[3]: 385–398. CrossRef.

McWilliams J (2014) The ecological creed of craft beer. *Conservation*, http://conservation-magazine.org/2014/03/sustainable-practices-in-craft-brewing/

Mintel (2013) The rise of craft beer in the US – craft beer sales have doubled in the past six years and are set to triple by 2017. January 23. http://www.mintel.com/press-centre/food-and-drink/the-rise-of-craft-beer-in-the-us-craft-beer-sales-have-doubled-in-the-past-six-years-and-are-set-to-triple-by-2017

Monloss K (2016) How Budweiser ditched its old look and crafted a new visual identity. *Adweek*, January 24. http://www.adweek.com/news/advertising-branding/how-budweiser-ditched-its-old-look-and-crafted-new-visual-identity-169144

National Gardening Association (2014) *Garden to Table: A 5-year Look at Food Gardening in America*. National Gardening Association, Williston, VT

Nurin T (2016) It's final: AB InBev closes on deal to buy SABMiller. *Forbes*, October 10. http://www.forbes.com/sites/taranurin/2016/10/10/its-final-ab-inbev-closes-on-deal-to-buy-sabmiller/#3155e6f137d6

O hUallachain B, Reid N (1991) The location and growth of business and professional services in American metropolitan areas, 1976-1986. *Annals of the Association of American Geographers* 81[2]: 254–70. CrossRef.

O hUallachain B, Reid N (1993) The location of services in the urban hierarchy and regions of the United States. *Geographical Analysis* 25[3]: 252–267. CrossRef.

Ohio City Neighborhood (2016a) About the neighborhood. http://www.ohiocity.org/about

Ohio City Neighborhood (2016b) Demographics. http://www.ohiocity.org/demographics

Ohio City Neighborhood (2016c) History. http://www.ohiocity.org/history

Ohio City Neighborhood (2016d) Neighborhood guide. http://www.ohiocity.org/guide

Pew Research Center (2010) Millennials: Confident. Connected. Open to change. Pew Research Center, Washington, D.C.

Porter M (1990) *The Competitive Advantage of Nations*. Basic Books, New York, NY. CrossRef.

Reid N, Gatrell JD (2015) Brewing growth: Regional craft breweries and emerging economic development opportunities. *Economic Development Journal* 14[4]: 4–12

Rothgerber H (2016) Better beer through Bellarmine chemistry. *Bellarmine Magazine*, Winter. http://www.bellarminemagazine.com/better-beer-through-bellarmine-chemistry/

Rotunno T (2014) The BMW of beer: Brewery crafts luxury approach. *CNBC*, May 3. http://www.cnbc.com/id/101636905

Satran J (2014) Here's how a six-pack of craft beer ends up costing $12. *HuffPost Taste*, September 12. http://www.huffingtonpost.com/2014/09/12/craft-beer-expensive-cost_n_5670015.html

Saxenian A (1994) *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press, Cambridge, MA

Schaeffer T (2016) Personal interview. Toledo, OH, February 9

Schönberger C, Kostelecky T (2011) 125th anniversary review: The role of hops in brewing. *Journal of the Institute of Brewing* 117[3]: 259–267. CrossRef.

Schnell SM, Reese JE (2003) Microbreweries as tools of local identity. *Journal of Cultural Geography* 21[1]: 45–69. CrossRef.

Schultz K (2015) Drink your way to sustainability with these climate-conscious microbrews. *Yes! Magazine*, April 3. http://www.yesmagazine.org/planet/drink-your-way-to-sustainability-with-these-climate-conscious-microbrews

Schumpeter JA (1942) *Capitalism, Socialism and Democracy*. Harper & Row, New York. CrossRef.

Scott A (1988) *Metropolis*. University of California Press, Berkeley, CA. CrossRef.

Scott A (1993) The new southern California economy: Pathways to industrial resurgence. *Economic Development Quarterly* 7[3]: 296–309. CrossRef.

Scott A (2006) Entrepreneurship, innovation, and industrial development: Geography and the creative field revisited. *Small Business Economics* 26[1]: 1–24. CrossRef.

Shilton AC (2016) Craft beer vs. Budweiser: How small brewers are winning back the neighborhood. *Yes! Magazine*, March 2. http://www.yesmagazine.org/new-economy/craft-beer-vs-budweiser-how-small-brewers-are-winning-back-the-neighborhood-20160302

Shinkle D (2016) Personal interview. Jeffersontown, KY, February 27.

Shortridge JR (1996) Keeping tabs on Kansas: Reflections of regionally based field study. *Journal of Cultural Geography* 16[1]: 5–16. CrossRef.

Smith N (1984) *Uneven Development*. Blackwell, Cambridge, MA. CrossRef.

Statistical Atlas (2015) Race and ethnicity in Lodo, Denver, Colorado. Cedar Lake Ventures, Inc. http://statisticalatlas.com/neighborhood/Colorado/Denver/Lodo/Race-and-Ethnicity

Storper M, Walker R (1989) *The Capitalist Imperative*. Blackwell, Cambridge, MA

Svaldi A (2016) Molson Coors claims all of MillerCoors, making home-grown brewer the third largest in the world. *The Denver Post*, October 11. http://www.denverpost.com/-2016/10/11/molson-coors-claims-millercoors/

The Beer Professor (2015) Community space. December 29. http://www.thebeerprofessor.com/?.p=1573

Thurnell-Read T (2014) Craft, tangibility and affect at work in the microbrewery. *Emotion, Space, and Society* 13[1]: 46–54. CrossRef.

Thurow LC (1978) Psychic income: Useful or useless? *The American Economic Review* 68[2]: 142–145

Treanor C, Costello B (2015) The Toledo metropolitan statistical area. *Economic Trends* September 3. https://www.clevelandfed.org/newsroom-and-events/publications/economic-trends/2015-economic-trends/et-20150903-the-toledo-metropolitan-statistical-area.aspx

Tremblay VJ, Tremblay CH (2009) *The U.S. Brewing Industry*. The MIT Press, Cambridge, MA

Törnqvist G (2011) *The Geography of Creativity*. Edward Elgar, Northampton. CrossRef.

Trubek A, Bowen S (2008) Creating the taste of place in the United States: can we learn from the French? *GeoJournal* 73[1]: 23–30. CrossRef.

Vandenengel H (2014) The fresh flavors of wet hop ales. *All About Beer Magazine* 35[5]. http://allaboutbeer.com/article/wet-hop-beer/

Walker E, Brown A (2004) What success factors are important to small business owners? *International Small Business Journal* 22[6]: 577–594. CrossRef.

Weilar S (2000) Pioneers and settlers in Lo-Do Denver: Private risk and public benefits in urban redevelopment. *Urban Studies* 37[1]: 167–179. CrossRef.

Wesson T, Nieva de Figueiredo J (2001) The importance of focus to market: A study of microbrewery performance. *Journal of Business Venturing* 16[4]: 377–403. CrossRef.

Winograd M, Hais M (2014) How millennials could upend Wall Street and corporate America. *Governance Studies at Brookings*, May. https://www.brookings.edu/wp-content/uploads/2016/06/Brookings_Winogradfinal.pdf

REGION

The Journal of ERSA
Powered by WU

Volume 4, Number 1, 2017, 51-70
DOI: 10.18335/region.v4i1.118

journal homepage: region.ersa.org

Critical Performance Factors for Large World Cities – In Search of Qualitative Causal Patterns by means of Rough Set Analysis*

Karima Kourtit[1], Peter Nijkamp[2]

[1] KTH Royal Institute of Technology, Stockholm, Sweden; Adam Mickiewicz University, Poznan, Poland (email: karima.kourtit@abe.kth.se)
[2] Tinbergen Institute, Amsterdam, The Netherlands; Adam Mickiewicz University, Poznan, Poland; KTH Royal Institute of Technology, Stockholm, Sweden (email: pnijkamp1@gmail.com)

Received: 9 January 2016/Accepted: 8 March 2017

Abstract. In the currently emerging '*urban century*', large cities and urban agglomerations are increasingly turning into (socio-)economic and political powerhouses of world-wide importance. This paper offers a comparative assessment of the multidimensional performance profile of major cities in our world, with particular emphasis on the livability and economic potential of these cities. We use an explanatory cohesive framework in the present study, which finds its origin in the so-called '*urban piazza*' model. Next, using an extensive database on various rankings of 40 global cities (the so-called GPCI data base), the relative performance of these cities is 'explained' in a qualitative sense by means of a non-parametric ordinal ranking method, known as rough set analysis. The approach allows us to identify in a qualitative sense the driving forces that altogether are responsible for the economic and livability achievements of these cities. Our analysis is based on a novel triple-layer rough set analysis of the performance of the cities concerned. The empirical analysis highlights the combined explanatory 'power' of several success conditions and drivers of the cities concerned. Our empirical results demonstrate that in particular geographical accessibility and R&D of these urban areas appears to play in many cases an important role in shaping their strong position.

Key words: urbanization; cities; performance; rough set analysis; piazza model

1 The Urban Planet

Demographic changes will be decisive for the future geography of our world (Tellier 2007). According to UN projections, our planet will house approx. 9 bln people by the year 2050. The rise with some 2.5 to 3.0 bln humans in the next 35 years – a time span covering a little bit more than one generation – is however, not equally spread over the world. There are significant regional differences, with a massive expected population rise in Africa, Asia and Latin America.

*This paper was inspired by several studies that have been produced in the context of the Joint Programming Initiative '*Urban Europe*' (see also Arribas-Bel et al. 2013). It is devoted to one of the icons in regional science, Manfred Fischer.

Another inevitable demographic megatrend for the decades to come is the massive population drift to the city. Cities will grow bigger – or rise in numbers – not only due to the global population rise, but also as a result of movements of people from peripheral and rural areas to urban agglomerations. There is a general consensus that the rate of urbanization is growing faster than the global population rise, so that our planet will become urban in nature[1] (see for more details also Kourtit 2014, 2015).

The above sketched rapid urbanization development will take place in particular in the developing world, so that growth of urban agglomerations will become a dominant characteristic of settlement patterns in the Third World. Other regions will be less affected by the disproportionally fast urbanization, while some regions (e.g., in Japan, Germany, Italy, Poland) may even face situations of shrinking cities (see also Kourtit 2015).

This study aims to provide a new urban analytics for tracing the determinants of the economic (market attractiveness, economic vitality, business environment) and livability (working, cost of living, security, life support functions) achievement of large cities in our world, based on a 'big data' system for these cities. This urban performance analysis is undertaken at the level of both aggregate and fine-gained data on these cities.

Large cities are, in general, powerful and effective growth engines in a modern global world. The rising importance of urban agglomerations emerges from so-called agglomeration externalities, partly in the form of MAR (Marshall-Arrow-Romer) externalities, and partly in the form of social network (or Jacobs 'melting pot') externalities of urbanization. Nonetheless, the relative efficiency of cities (in terms of productivity, resource use etc.) may differ significantly, and hence it is a matter of empirical research to trace which city has the highest performance and under which conditions (see also Kourtit 2014, 2015).

The emerging new urbanization patterns that manifest themselves world-wide prompt an avalanche of new policy and research challenges and concerns. Particularly in such domains as income distribution, climate change, quality of life, food, energy, safety, mobility, poverty and social cohesion, major problems in cities are expected. In the developing world, it is foreseen that housing problems – in particular, slums – will continue to exist and will have a detrimental effect on the daily well-being of a significant share of the world population. Even more severe problems are anticipated for the decades to come: sea level rise, floods, social tensions, refugee flows, extreme weather conditions, traffic deadlocks, and so forth. It goes without saying that ideally urban economic development has to proceed in tandem with urban ecological sustainability. At present, more than 75 percent of the global environmental impacts are produced by just over 50 percent of the world population, leaving behind a threatening ecological footprint. Is economic growth and productivity able to cope with these issues? Such major and world-wide concerns call for strict governance systems and response mechanisms, but this

[1] In a standard textbook on urban economics by Mills, Hamilton (1994), the authors raise the question: 'What are urban areas?'. They argue that there are many urban concepts: town, city, urban area, metropolitan area, and the like. Some of these concepts are based on legal definitions, and hence these definitions vary across nations. Others are based on population figures and reflect the fact that in urban areas the average population density is higher than elsewhere. The authors provide the following basic description: 'Thus the fundamental and generic definition of an *urban area*, or *metropolitan area*, is a place with a much higher population density than elsewhere' (p. 3). It is thus apparent that the concept 'urban' is not a clearly operational term. In our study, we use this term as a general description of a place or area with a relatively high concentration of people in a demarcated built environment. To be slightly more precise, in the present study, we adopt the following nomenclature in the present paper (see the *Dictionary of Human Geography*, edited by Gregory et al. (2009)):

- *urbanization*: has often been understood as a process of human distancing from first nature. In cities, nature became a residual or artificial category limited to parks, zoos and urban – mostly ornamental – gardens.
- *city*: today, a more generic usage of this term refers to an urban demographic, economic and above all political and jurisdictional unit, usually bigger than a town.
- *agglomeration*: the association of productive activities in close proximity to one another. It typically gives rise to external economies associated with the collective use of the infrastructure of transportation, communication facilities and other services. Agglomeration also facilitates the rapid circulation of capital, commodities and labour.
- *urban areas*: may be cities or towns and are characterized by a higher population density in comparison with the areas surrounding them.

is exactly the Achilles heel of modern urban planning and policy-making. Clearly, we also need reliable and comparable databases on urban development and their backgrounds, in order to develop effective policy intervention tools (Batty 2013, Vaz 2016, Vaz, Arsanjani 2015, Vaz et al. 2015).

Fortunately, we observe nowadays a renewed and world-wide interest in urban challenges and responses, as well as in the strategic position of world cities in global networks (see Neal 2012, Taylor 2004). Major issues in this debate are:

- The impact of population dynamics on urbanization patterns (in particular, cross-border migration movements).
- The relationship between economic cycles and urban developments (including shrinking cities, such as Detroit or Leipzig).
- The acceleration in urbanization rates in the developing world, and the far-reaching consequences for the economy, society and ecology.
- The serious risks encountered by massive population concentrations, which may be subjected to major catastrophes (e.g., floods, social unrest, extreme weather conditions, terrorism, power cuts).
- The persistent – and sometimes sharpening – poverty traps in many large agglomerations in our world.
- The logistic control of extremely complex urban physical and virtual networks (e.g., commodity distribution, mass transit, volatility of geographic locations, digital data management for public policy etc.)
- The span of control of modern urban governance systems in the management and strategy development of the rising number of mega-cities on our planet.

It is clear that modern large cities are complex spatial entities, with a great heterogeneity in socio-economic, transport, business and ecological outcomes. Several cities appear to have a much better performance net outcome than others. And therefore, it is important to trace the drivers of the achievement levels of large cities in our world. It is clear that economic outcomes need to be explained from a broader perspective, while also the quality of life plays an important role. The theoretical underpinning of the design of our research can be found in the 'urban piazza' model in Figure 1, to be highlighted in Section 2.

Clearly, the great challenge for large cities in our world is to perform much better. Many cities in our world face sometimes similar, sometimes different problems in achieving a respected and recognized international profile. We talk more and more about global urban competition and competitiveness profiles of cities. This so-called '*rat race*' (Kourtit, Nijkamp 2013b, Kourtit, Nijkamp, Suzuki 2013) does not mean that cities are in all respects competitors which operate on a 'predator-prey' basis, but it suggests that many world cities aim to reach a top position in a global competitive ranking system. The aim of the present paper is now to offer a methodology for comparing in a logical and systematic way a broad range of qualitative performance[2] indicators of a set of global cities, in total 40 in number. This comparative study uses a very extensive and rich database from the Global Power City Index system developed by the Japanese Mori Memorial Foundation (2012). This dataset comprises detailed information on numerous indicators for the cities under consideration. In the annual GPCI database reports, a qualitative ranking of these 40 cities – on the basis of their multidimensional performance profile – can be found. The challenging research question in this paper is now whether in a systematic way the determinants for the performance ranking of these cities can

[2]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 economic achievement position of an actor or corporate organization. The latter meaning will also be adopted in our study on global large cities.

be identified. Given the qualitative rank-order information on the determinants, a combinatorial deterministic method will be used, so that factors that are of critical importance (i.e., necessary conditions) for the economic and ecological performance of cities can be traced. The comparative assessment methodology employed here originates from an artificial intelligence method, coined rough set analysis, which is very appropriate in assessing the critical success conditions (stimuli) that lead to a set of output indices (responses) from a number of competitors or alternative choice possibilities. This method which finds its origin essentially in ordinal multi-criteria analysis will concisely be presented in Subsection 2.2.

The paper is organized as follows. Section 2 will be devoted to the description of the framing and the database of our study. Then in Section 3, the operational methodology employed will be briefly outlined, and will offer an overview and interpretation of the empirical results. Section 4 will then conclude with some retrospective and prospective remarks.

2 Framing of the Research

2.1 Description of the database

The rising interest in city growth and size and in urban networks has in recent years prompted a wide variety of city concepts, such as smart city, global city, world city, mega-city, metropolis, ecumenopolis, megalopolis, connected city etc. (Doel, Hubard 2002, Kourtit, Nijkamp 2013b, Neal 2012, Sassen 1991, Taylor 2004). They all have different meanings, but are also often used interchangeably in the literature. In the present paper we will use the neutral term large city, which may in general means that the city concerned – or the urban agglomeration – will have more than 1 mln inhabitants. Our interest is not in conceptual purity, but in the analytical question: what makes the difference? In other words, if a given large city can be characterized by a set of manifold performance indicators, which are the conditioning variables (critical success factors) that determine in a relative sense the success of the city at hand, and to what extent?

Clearly, in principle, the number of performance criteria, as well as the number of drivers (inputs or stimuli) is vast, while sometimes the distinction between drivers and output is not clear at all. Very often, economic characteristics – for instance, number of headquarters of international banks, financial services, international events or exhibitions, or economic image – are seen as performance variables (outputs), while public investments – e.g., international transport infrastructure, advanced IT networks, health care facilities – are often regarded as support conditions that may act as drivers (or inputs). But in a way, they might also partly be seen as endogenous success indicators. Thus, there will always be some ambiguity in the definition of drivers and responses (Kourtit 2015).

In our applied analysis, we will use the extensive and unique multi-dimensional database developed by the Institute for Urban Strategies (Mori Memorial Foundation 2012), which contains a systematically collected and annually updated vast amount of interesting urban data (starting from 2008) that are relevant for our purposes. The data base comprises 40 large cities, with a fair distribution between developing and developed countries. The data is a mix of survey data and ordinal response and perception data. This data base is subdivided into six main categories: *Economy*, *Research & Development*, *Cultural Interactions*, *Livability*, *Environment*, and *Accessibility*. These six categories in the database are next decomposed into a total of 69 empirically tested and mutually comparable sub-indicators for each of the 40 large cities in the world that are considered. This list of preselected cities is a priori given for our analysis; see Figure 2 for this complete list. A very interesting characteristic is also that the database considers the classes of relevant global and local actor-specific scores and rankings of the socio-economic performance of cities from the perspective of prominent stakeholders, in particular, residents, visitors, artists, researchers, and managers (see also Arribas-Bel et al. 2013, p. 252). This is thus a rich and original database, which is particularly suitable for applied comparative research on urban performance.

The GPCI database is based on a cross-section of identifiers of urban development, based on interviews and data collections from local stakeholders. Clearly, each city may have its own desired performance profile with unique or place-specific indicators, but in an internationally comparative and competitive benchmark exercise cities have to be compared on common and identical indicators. As in a global force field it is the external world which decides on the successfulness of urban agglomerations.

This Global Power City Index (GPCI) database offers a multidimensional picture of the socio-economic performance and position of different global cities in the world from the perspective of attracting and retaining firms, talent and investment to these cities on the basis of inter alia quality of life, culture and accessibility, complemented with detailed information on the perceptions of relevant groups of stakeholders (see also Tranos et al. 2014). Table 1 offers a brief overview of the main classes of performance indicators derived from the GPCI-2012 data system (see also Arribas-Bel et al. 2013, p. 252). This information is used in the GPCI analysis to create a performance ranking of all 40 cities. The main question is now: What are the drivers of this ranking?

In highlighting the rank orders of the individual urban indicators included in the six main categories used in the GPCI-2012, the information base demonstrates the different strengths of the socio-economic achievement levels of the leading global cities. This rich information base can also help to improve our understanding of the attractiveness of a particular large city (see also Arribas-Bel et al. 2013, p. 252), for instance, through a benchmark analysis. These GPCI-data are updated annually in a careful and reliable way, so that in principle not only a cross-sectional comparison of cities can be made, but also a combined space-time panel assessment. In the present study however, we will limit ourselves to a cross-section assessment of the performance of cities, mainly since the relative changes in urban performance on an annual basis appear to be rather modest.

Table 1: List of performance indicators of the GPCI-2012

Indicator Group	Indicator
	ECONOMY
Market attractiveness	1 GDP 2 GDP per capita 3 GDP growth rate
Economic vitality	4 Total market value of listed shares on stock exchanges 5 Number of world's top 300 companies 6 Number of employees
Business environment	7 Unemployment rate 8 Number of employees in service industry for enterprises 9 Average wage level (compared to New York) 10 Easiness of securing human resource 11 Office area per employee
Regulations and risks	12 Index of economic freedom 13 Corporate tax rate 14 Index of country risk (political, economic, business, etc.)
	RESEARCH & DEVELOPMENT
Research background	15 Number of researchers 16 World's top 200 universities 17 Basic skill of mathematics and science
Readiness for accepting and supporting researchers	18 Readiness for accepting foreign researchers 19 R&D expenditure
Achievement	20 Number of registered industrial property rights (patents) 21 Number of highly-reputed prize winner 22 Activeness of interaction between researchers and outputs of their achievement
Trendsetting potential	23 Trade value of audiovisual and related services 24 Number of holdings of international conventions 25 Number of holdings of world-class largest cultural events 26 Environment of creative activities

Continued on next page

Table 1 – continued from previous page

Indicator Group	Indicator
Accommodation environment	27 Number of guest rooms of luxury hotels 28 Number of hotels
Resource of attracting visitors	29 Number of world heritages (within 100km area) 30 Cultural attractiveness, etc. 31 Number of theatres and concert halls 32 Number of major museums 33 Number of stadiums
Shopping & Dining	34 Satisfactory level of shopping 35 Satisfactory level of dining
Volume of interaction	36 Number of foreigners 37 Number of visitors from abroad 38 Number of foreign students
	LIVABILITY
Working environment	39 Total working hours 40 Satisfactory level of employees' life from the viewpoint of managers
Cost of living	41 Average rent (residential) 42 Average price level (compared to New York)
Security and safety	43 Number of murders 44 Vulnerability 45 Healthy Life Expectancy 46 Activeness of community
Life support functions	47 Population density 48 Number of medical doctors per residents 49 Number of international schools per foreign residents 50 Variety of retail shops 51 Variety of restaurants
	ECOLOGY AND NATURAL ENVIRONMENT
Ecology	52 Number of companies with ISO 14001certification 53 Percentage of renewable energy 54 Percentage of recycling 55 CO2 emissions
Pollution degree	56 Density of suspended particulate matter (SPM) 57 Density of sulfur dioxide (SO2) 58 Density of nitrogen dioxide (NO2) 59 Water quality
Natural environment	60 Situation of green coverage 61 Average yearly temperature differences
	ACCESSIBILITY
Infrastructure of transportation	62 Travel time between inner-city areas and int'l airports 63 Number of cities with international direct flights 64 Number of travellers of international flights 65 Number of runways
Infrastructure of inner-city transportation	66 Number of stations (subway) 67 Punctuality of public transportation (train, subway, bus) 68 Satisfactory level of commuting 69 Taxi fare

Source: Mori Memorial Foundation (2012)

The data in the GPCI-database have a broad coverage. But they have to be treated in a logical and coherent framework, so as to identify in a systematic and consistent way the determinants and responses of urban systems in a comparative sense. All details on the collection, standardization, definition and geographical scale of the data can be found in the above mentioned GPCI-report. Annex A contains more detailed rankings of these cities on the basis of the above mentioned principles (Arribas-Bel et al. 2013, p. 254).

In explaining in a qualitative sense the rankings of the cities at hand on the basis of their performance profile, we have focused on two categories of responses or achievements, viz. the economic performance and the livability in each of these cities. Consequently, these two categories of endogenous variables are assumed to be determined by a wide range of background variables, so that all remaining indicators are assumed to be driving forces of these achievement rankings of these cities. This will be tested in our empirical work.

The methodological and conceptual framework of our research on the performance of 40 global cities originates from the '*urban piazza*' architecture (see Figure 1). The '*urban piazza*' architecture acts here as the integrated conceptual framework and navigation tool for assessing performance criteria of the global cities under varying condition variables (or stimuli). An extensive underpinning and treatment of the 'urban piazza' concept can be found in Kourtit (2015); we refer here also to Arribas-Bel et al. (2013) and Kourtit et al. (2014). The essence of the piazza model is that it offers a comprehensive ideal-typical representation of the urban force field that generates the performance of a given city.

Historically, the 'piazza model' is essentially based on the general functional coherence of and socio-economic independencies of activities in ancient Italian cities, in which all relevant forces (economic, social, political, logistic, etc.) were coming together in a central place or node which generated the heartbeat of the city concerned. It offers a broad framing of urban functionalities and mechanisms, from which we have selected a limited set of focused variables that are in agreement with the GPCI dataset. In our study, we assume that nowadays four such major drivers can be distinguished, which are of critical importance for two classes of response (or output) variables. The piazza architecture used in the present empirical analysis is therefore, decomposed into four main segments viz., *R&D*, *Cultural Interaction*, *Environment* and *Accessibility*, which form the four major condition variables to explain separately two major endogenous decision variables (or responses), viz., XXQ *Economic Performance* and XXQ *Livability* (see Figure 1). In other words, global cities seek to generate the highest possible productivity and quality of life (similar to the general XXQ-concept, introduced by Nijkamp (2008)) so as to attract and retain creative and talented actors (employees, self-employed people or entrepreneurs, but also residents and visitors) in order to generate positive externalities (Kourtit, Nijkamp 2013a,c).

Thus, the above two XXQ indicators are a measure of global cities' performance in which inputs or efforts (Cultural Interaction, R&D, Accessibility and Environment) are linked to achievements of the city at hand (e.g., profitability, vitality, sustainability, etc.). Clearly, the piazza concept can be positioned in a broader context of recent spatial growth models (including absorptive capacity) and urban creativity analysis. It is evident that the presence and experience (individual values, preferences and visions) of creative minds can create critical conditions for enhancing the level of attractiveness of cities, leading to a concentration of innovative potential in geographical space (Kourtit, Nijkamp 2013a,b,c,d, Kourtit 2015). The piazza framework will now be used as an integrating basis for a broad urban assessment exercise, with the aim of tracing the roots of the economic success (XXQ *Economy*) and livability (XXQ *Livability*) of cities on the basis of their innovative capacity (*R&D*), cultural ambiance (*Cultural Interaction*), natural and ecological quality (*Environment*) and transport access and connectivity (*Accessibility*). The information contained in Table 1 is partly quantitative, partly qualitative in nature. To avoid an ambiguous analysis of heterogeneous data, all information was consistently translated into rank-order data.

Since the rankings of the 40 cities under investigation are based on multivariate data that are combined to create a qualitative rank order, it is necessary to seek for an analytical explanatory model that is able to encapsulate such qualitative data. Consequently, we need to look for a qualitative explanatory model, known as *rough set analysis*. This will be outlined in the next subsection.

2.2 Rough set analysis as a research tool

For the analytical part of our qualitative comparative assessment study, a multivariate qualitative classification method, called rough set analysis (RSA), is used. This method

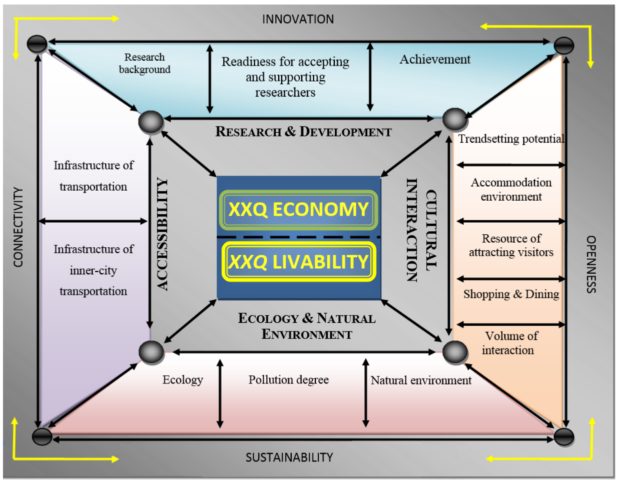


Figure 1: The '*urban piazza*' for the performance structure of large cities

is inspired by the artificial intelligence methodology (for more details, see also Pawlak 1991, Pawlak, Slowinski 1996, van den Bergh et al. 1998, Slowinski, Stefanowski 1994, Greco et al. 1997). This approach will be used as the main methodology, in order to offer an explanatory non-parametric and qualitative model for assessing the critical success conditions that lead to a set of output indicators (responses) for each of the rivalry choice options. This rough set method will be applied to operationalize the conceptual piazza model depicted in Figure 1.

RSA is a deterministic method, which is based on combinatorial set theory (see also Baaijens, Nijkamp 1997, p. 11)(see also Baaijens and Nijkamp 1997, p.11), and may be regarded as a qualitative explanatory regression analysis in case of small samples. A disadvantage of a standard regression analysis is that, if the sample size of the model is limited, the reliability of the regression results is rapidly going down, especially in case of a large number of explanatory variables. In addition, in case the relevant variables have only a qualitative or nominal meaning (e.g., tourism loyalty, historical heritage), application of quantitative regression techniques becomes problematic. RSA is particularly suitable in case of qualitative information and small samples (see Kourtit, Nijkamp, van Leeuwen 2013). The main challenge of RSA is to provide an analytical approximation of logically consistent and permissible statements from the acquired available database. It takes for granted the general scope of regression analysis, viz. that a set of condition variables (explanatory variables, stimuli) impact a set of endogenous variables (responses or decision variables). The main idea is then to find out under which conditions the stimuli have a decisive impact on the responses. In contrast to regression analysis where impact coefficients are estimated, rough set analysis – based on qualitative information – then tries to identify the frequencies of occurrences from the data set, with which a set of stimuli exerts an influence on one or more response variables. This means essentially a deterministic analytics in tracing the impact of condition variables on endogenous variables. From this perspective, RSA is much richer in scope than simple correlation or regression analysis, also because it is able to consider simultaneously more than one endogenous variable, as is the case in one study.

The rough set method is thus suitable to deal with compound qualitative information, and offers a consistent description of a set of qualitative variables on the basis of lower and upper approximations in case numerical values of these variables do not exist. This information is considered as a finite set of objects, which can be described through a set of characteristic attributes. These attributes are supposed to construct the available information on the objects (alternatives, items). On the basis of a set of attributes, a rough set approach can categorize phenomena into groups with identical features by considering relevant measurements that may weaken the 'unobserved heterogeneity'.

RSA is also an original method for dealing with incomplete information patterns, and may therefore be a useful approach for a transversal comparison of urban achievements from a broad multidimensional perspective (see Orlowska 2013). Data reduction is often the main element of RSA, as it allows to identify and represent hidden structures in the data (see also Baaijens, Nijkamp 1997, p. 11). The result of a RSA of a multidimensional data set is ultimately a decision table which forms the basis for deriving deterministic decision rules, which are compatible and logically consistent with the qualitative pattern of the data. The decision rules are normally of a conditional (i.e. '*if* ... *then* ...') nature and indicate also the strength of such conditional statements (i.e., the number of objects – or frequencies – for which this statement is true) (see Kourtit, Nijkamp, van Leeuwen 2013)(see Kourtit et al. 2013b). These expressions describe the associations between the objects considered and their assignment to particular classes (see Pawlak 1991, 1992)(see Pawlak 1991; 1992). We will not provide here an extensive discussion of the many set-theoretical technalities of RSA. Detailed information on RSA both from a conceptual and an applied perspective can be found in Pawlak (1991), Fayyad et al. (1996), Deogun et al. (1997), Famili et al. (1997), Slowinski (1995), van den Bergh et al. (1998), Ziarco (1998), and Wu et al. (2004).

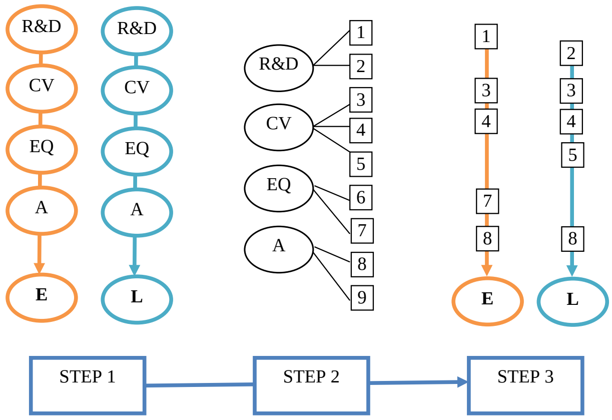
RSA can nowadays be applied by using standardized open access software. The RSA in the present study is performed with the support of the software programme Rough Set Data Explorer (*ROSE*). This programme is instrumental to calculate the basic minimal covering of conditional statements from a complex dataset.

We will now use the *ROSE* software here to compare, in a logical and systematic way, the performance indicators of the set of 40 global cities (i.e. the number of objects) from the GPCI data base in terms of 2 output variables (economy, livability) and 4 input variables (accessibility, R&D, cultural interaction, ecological & natural environment) (the number of attributes considered is thus 6), as outlined in Table 1 and Figure 1. In this way, rough set theory helps to provide insight into the multiple achievement categories of objects and to highlight the most prominent factors behind their achievements and performances (see also Baaijens, Nijkamp 1997). RSA provides thus an operational multidimensional classification method that helps identify important critical factors for a systematic performance analysis of competing objects. In the next section, we will present the empirical findings of the rough set analysis applied to the above mentioned GPCI data set.

3 Rough Set Analysis of Composite and Single Drivers of Urban Performance

Following the conceptual model of Figure 1, we have – as mentioned above – chosen to focus on qualitative *XXQ Economy* and *XXQ Livability* indicators for the 40 cities under research. The reason for this choice is that these factors are both distinct from each other and can be well explained from a causality perspective by the other four (composite) factors (R&D, Cultural values, Environmental quality and Accessibility). Thus, the aim is to develop a qualitative explanatory model for the case of two types of performance indicators (decision variables) of GPCI cities as a result of the existence and functioning of four types of underlying drivers (condition variables, in rough set terminology). Thus, in fact, the urban piazza framework of Figure 1 is translated into a qualitative causal model explaining (a set of) two endogenous output variables from (a set of) four exogenous explanatory (input) variables. These output and input variables are, in the RSA jargon, often called decision and condition variables, respectively.

Our RSA will be applied in three steps, as shown in Figure 2. In the *first* step, we only use the main composite variables, both as decision (2) and condition (4) variables, to test globally the causal pattern implied by our conceptual model from Figure 1. In the *second* step, we will have a closer look at the distinct sub-factors from the 4 condition variables which are the most influential ones for each of the two overall composite factor scores. Finally, in the *third* step, we will use the most important critical sub-factors that resulted from step 1 and 2 to explain the composite performance factors *XXQ Economy* and *XXQ Livability* as decision variables.



Notes: E = XXQ Economy; L = XXQ Livability; R&D = Research & Development; CV = Cultural Interaction; EQ = Ecology & Natural Environment; A = Accessibility

Figure 2: Three steps in the rough set approach using composite factors and sub-factors

Table 2: Rough set rules explaining *XXQ Economy* by composite factors using 4 classes

Rule #	R&D	Cultural	Environment	Accessibility	Econ(D)	Coverage (%)
1	1				**1**	42
2				1	**1**	29
3			4	4	**1**	29
4		1	3		**2**	22
5			3	2	**2**	29
6	2	1	1		**2**	21
7	3		4		**3**	42
8	2	1	4		**3**	25

3.1 Explaining composite urban performance factors from composite driving forces (Step 1)

In the first step of our analysis we will explore how the 6 composite factors are related to each other. Therefore, the composite economic indicator from Table 1 is used as the decision variable (in rough set terms), to be explained by 4 background variables or condition variables. We classify now the 40 cities from the GPCI data base first into four groups, based on their *XXQ Economy* indicator. The rankings are denoted as 4 qualitative classes of features, viz. low, medium-low, medium-high and high. The RSA results show that for the *XXQ Economy* and *XXQ Livability* scores, all 4 conditional variables appear to be in the 'core' of the rough set classification. This means that each of them is relevant and necessary in explaining the level of *XXQ Economy*. Next, a sensitivity analysis shows that this finding also holds when distinguishing three classes of – rather than four classes – *XXQ Economy*, so that this result is rather robust.

The quality of classification (viz., in rough set terms, the significance of the combination of background variables playing an explanatory role) appears to be quite high, namely 0.80. With these settings, 8 so-called decision rules or logical statements that are consistent with the data can be derived from the rough set analysis (see Table 2).

The strongest decision rules appear to be rules 1 and 7, which both cover 42% of the dataset. These tell us that a low R&D level results in low economic development; and a medium-low level of R&D combined with a high level of environmental quality results on average in a medium-low economic score for the cities under consideration. There is apparently quite some degree of heterogeneity among the 40 cities in our sample, so that each separate statement from the total set of 8 all rules is somewhat difficult to interpret.

When we apply next a sensitivity analysis and split the cities into 3 groups, with the categories low, medium and high *XXQ Economy*, again all condition variables appear to

Table 3: Rough set rules explaining *XXQ Economy* by composite factors using 3 classes

Rule #	R&D	Cultural	Environment	Accessibility	Econ(D)	Coverage (%)
1	1		1		**1**	77
2	1			2	**1**	46
3		2	3	2	**2**	27
4	3				**3**	33
5		3			**3**	25

Table 4: Rough set rules explaining *XXQ Livability* by composite factors using 4 classes

Rule #	R&D	Cultural	Environment	Accessibility	Livability(D)	Coverage (%)
1		3			**1**	23
2	3		2		**1**	23
3	3			1	**1**	23
4	1			1	**1**	23
5			2	1	**1**	23
6		2			**2**	33
7		4		2	**2**	33
8			1	2	**2**	33
9	3	4	3		**2**	33
10	3		3	4	**2**	33
11			4	4	**3**	27
12			2	2	**4**	20
13	2	4	4		**4**	30
14	2	4		3	**4**	30
15	2		4	3	**4**	30

be in the core. However, the quality of the classification is slightly lower than before, viz. 0.68 for the economic system (Table 3).

Table 3 shows that the first two rules are the strongest ones, with a coverage of respectively 77 and 46 percent. It is clear that a low R&D, even with a medium accessibility, is related to a low economic performance. Apparently, R&D is a *sine non qua* for the economic achievements of cities. The other three rules all have a coverage above 25 percent. These indicate that a high level of R&D and a high level of cultural quality are related to high levels of economic performance.

Next, when we take a look at the *XXQ Livability* scores, again all condition variables appear to be in the core, with a classification quality of even 0.95. This results in a relatively large number of rules, again obtained through using a minimum of 20% of satisfaction (in rough set terms) and always a maximum number of 4 condition variables (see Table 4). Thus, all 4 explanatory factors may be seen as determinants of the urban output variables.

Next, when using again 3 classes of *XXQ Livability*, again all condition variables appear to be in the core; however, the quality of the classification is a little lower than before, namely 0.40. This also results in a lower number of satisfactory rules, namely only 6, of which 5 deal with low Livability scores. In addition, except for rule 5, all rules have a satisfaction level between 21 and 23%.

The clearest conclusion from the above results is that low levels of accessibility are related to low Livability scores. This holds for the strongest rule, number 5, with a satisfaction score of 38%, as well as for the 4 subsequent rules. Those first 4 rules indicate that even when levels of R&D or culture are high, but urban accessibility is low, Livability is low. Rule 6 indicates that a medium level of R&D with a medium level of culture is related to medium Livability scores. Clearly, these findings may be seen as rather plausible outcomes.

Table 5: Rough set rules explaining *XXQ Livability* by composite factors using 3 classes

Rule #	R&D	Cultural	Environment	Accessibility	Livability(D)	Coverage (%)
1	1				**1**	23
2	3			1	**1**	23
3		3		1	**1**	23
4		2		1	**1**	23
5			1	1	**1**	38
6	2	2			**2**	21

3.2 Analysis of sub-factors (Step 2)

The dataset we use is a very rich one. The 6 main indicators are grouped into various sub-factors (see Table 1), as are the sub-factors themselves. To see how these sub-sub factors relate to the main indicators, we perform in our rough set approach a heuristic search for reducts (a subset of attributes which can, by itself, fully characterize the information contained in the database) which, in other words, have a basic minimal covering (see also Pawlak 1991, Pawlak et al. 1995, Pawlak, Skowron 2007, Poel 1998, Polkowski 2003). Within the *ROSE* software program, attributes are added to the existing core by selecting the most promising indicators from this set and by trying different paths. A higher frequency rate means that the sub-factor is more important in explaining the level of the composite score. Appendix A shows the percentages of all sub-factors per composite factor.

First, when looking at the 13 sub-actors of the *Economy* factor, 5 of them are never part of a minimal reduct. This means that they are the least related to the overall score on the Economy factor. However, 3 of the sub-factors have a score of 100 percent, which means that they are part of all minimal reducts and thus of the core of the dataset. These are: GDP per capita, presence of world top 300 companies, and office area per employee.

Livability is also composed of 13 sub-factors. Of these, 3 have a frequency score of 100 percent. Among those are the number of murders and the variety of shops and restaurants. Apparently, cities with a high degree of criminality can still have a high level of Livability. Furthermore, we can distinguish 1 core variable, namely, maturity of the community. In addition, 2 other variables have a high frequency score: average house rent and medical doctors per inhabitant.

Research and development is composed of 8 sub-factors, which all appear at least once in a minimal reduct. Two of them appear in every reduct and have a 100 percent frequency score; these are: interaction opportunities between researchers, and the number of winners of highly-reputed prizes.

Cultural interaction is composed of the highest number of different sub-factors, viz. 16. Of those, only one half appears in one of the minimal reducts. Three of them are in the core of the dataset: number of theatres and concert halls, number of visitors from abroad, and opportunities of cultural, historical and traditional Interaction.

The next composite factor is *Environment*, which consists of 9 sub-factors of which as many as four are part of the core. These are: CO2 emissions, density of suspended particulate matter (SPM), level of green coverage, and comfort level of temperature.

The final composite factor is *Accessibility*. Of the 10 sub-factors, 8 appear once or more times in a minimal reduct. One of them is a core variable, which is punctuality of public transportation. Apparently, this variable is very important in explaining high levels of accessibility. Furthermore, density of metro stations and the presence of direct international freighter flights appears to be also important.

This heuristics search exercise results in an interesting list of 18 single sub-factors that can be used to explain the composite factors *XXQ Economy* and *XXQ Livability* (see Table 6). Since 18 condition variables is a high number, we thus perform another heuristic search on the list of most important single factors with either *Economy* or *Livability* as decision variables. We selected all variables with a minimal reduct score of 50% and above for the final rule induction (variables are indicated in bold).

Table 6: The most important single sub-factors

	XXQ Economic	XXQ Livability
R&D		
Number of Researchers	34	4
World's Top 200 Universities	53	61
Readiness for Accepting Foreign Researchers	25	100
Cultural		
Number of International Visitors	88	
Level of Satisfaction for Dining	25	9
Number of Museums	2	22
Number Hotels		35
Environment		
CO2 Emissions	35	17
Density of Suspended Particulate Matter (SPM)	31	4
Level of Green Coverage	51	65
Comfort Level of Temperature	29	
Number of Companies with ISO 14001 Certification	81	52
Density of Sulfur Dioxide (SO2), Density of Nitrogen Dioxide (NO2)	15	87
Percentage of Renewable Energy Used	41	61
Accessibility		
Density of Metro Stations	80	100
Travel Time between Inner-city Areas and International Airports		9
Number of Passengers on International Flights	22	17
Transportation Fatalities per Population	20	

3.3 Explaining Economy and Livability from the most important critical sub-factors (Step 3)

Finally, the third step of the analysis performed in this study uses the sub-factors that are part of at least one of the minimal reducts. We impose a minimal coverage of 25 percent and a maximum number of attributes of 3 items.

When looking at the single factors explaining *XXQ Economy*, there are some interesting findings. Clearly, the presence of good universities appears to be beneficial. A low score for this variable, together with medium/low green coverage and a low number of international visitors, is related to a low score on *XXQ Economy*. This holds, for example, for Bangkok, Cairo, Fukuoka, Moscow and Mumbai. At the same time, a high score on top-200 universities, together with a high score on certificates, is related to a well performing Economy. This holds for example for Beijing, Hong Kong, London and Tokyo.

In addition, the number of international visitors affects the quality of the economy: rule 6 indicates that cities like Beijing, London, New York, Paris, Sydney and Barcelona, with a high number of international visitors, have a high score on *XXQ Economy*.

When looking at Livability, it appears that the urban density of metro-stations has a positive effect. In Beijing, Geneva, Toronto and Zurich the density is low and the livability of medium quality, while in Barcelona, Copenhagen, New York, Paris and Tokyo the density is high, as is Livability. Also the readiness for foreign researchers seems to positively affect *XXQ Livability*. In Cairo, Moscow, Mumbai and Sao Paulo, for example, this readiness is low, as is the Livability. Green coverage and the presence of top universities shows mixed results. Our hypothesis that the economic and ecological performance of large cities is largely determined by a combination of 4 categories of empirical evidence on 40 large cities is confirmed in our analysis, be it that individual exceptions may occur (just like in any normal regression analysis).

Table 7: Rough set rules explaining *XXQ Economy* by the most important single subfactors

Rule	Top 200 Universities	Green coverage	Internatio-nal visitors	Certificates	XXQ Economy (D)	Coverage (%)
1	1	2			**1**	31
2	1	1	1		**1**	69
3	1	1		1	**1**	54
4			2	1	**2**	60
5		2	2		**2**	27
6			3		**3**	50
7	3			3	**3**	33

Table 8: Rough set rules explaining *XXQ Livability* by the most important single subfactors

Rule	Top 200 Univer-sities	Readiness for foreign researchers	Green coverage	Dioxide	Renew-able	Metro	XXQ Livability (D)	Coverage (%)
1		1					**1**	31
2			2				**1**	31
3	2					1	**2**	36
4					2	1	**2**	29
5	1	3					**2**	29
6		3	1	2			**2**	36
7						3	**3**	38

4 Concluding Remarks

In an increasingly urbanized world, cities – or in a more general sense, urban agglomerations – are becoming engines of growth and sustainability. In the emerging '*New Urban World*' (see Kourtit 2014, 2015), cities are not functioning as isolated islands, but have to compete in terms of their socio-economic and attractive achievements with other cities. Therefore, it is of utmost importance to identify the drivers of these achievements. In our study, we have presented the 'urban piazza' as a conceptual and coherent framework for mapping out the force field of modern cities, from the perspective of identifying and assessing the critical success factors for urban performance measured along two multidimensional criteria, viz. Economy and Livability.

Our study has used a comprehensive database for 40 cities all over the world to trace the drivers of urban performance, using the GPCI-database with qualitative rankings of all these cities. Our study has brought to light that there is not a 'one size fits all' result from our explanatory model. Different cities appear to have a varying performance through the influence of various important drivers. In many cases however, it turns out that accessibility and R&D may be seen as a major critical success condition, though in several cases in combination with other determinants of urban performance.

Urban performance analysis, as exemplified by our study, may become an important research tool in the future. Next to standard quantitative data, also qualitative information may prompt many insights into the success conditions of cities. Clearly, in the future, urban 'big data' may become another rich source of information, in order to better understand the fine-grained nature of urban dynamics and growth (Vaz 2016, Vaz, Arsanjani 2015, Vaz et al. 2014, 2015). It goes without saying that urban performance analysis may open new departures for research in the urban sciences. The rough set methodology applied here is a promising example of such a new analytical tool.

Clearly, RSA is able to act as an identifier in a complex pattern recognition exercise. Through a combinatorial algorithm, it offers a tool for conditional qualitative causal statements, in particular, by examining the frequency of validity of a statement on the presence of an explanatory variable impacting – either as a single variable or in combination with other variables — one or more endogenous variables. Clearly, there are three marked differences with a standard regression model: rough set analysis handles qualitative (even nominal) information; it is deterministic in nature; it can handle multiple response variables.

Clearly, the present study has some limitations. The sample of 40 cities is not very large in number, but this is largely compensated for by the in-depth systematic and accurately tested information provided on each individual city. Next, the benchmark picture of these 40 cities has a limited time horizon of at most 3 years, so that a really dynamic longitudinal perspective is lacking. Thus, there is certainly scope for analytical improvement in the future, so that in the end also more advanced urban econometric models might be developed and tested.

Acknowledgement

The authors wish to thank Eveline van Leeuwen for her great assistance in running the software for the rough set analysis.

References

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

Baaijens S, Nijkamp P (1997) Time pioneers and travel behaviour: An investigation into the viability of slow motion. Tinbergen Institute Discussion Papers 97-132/3, Tinbergen Institute, Amsterdam

Batty M (2013) *The New Science of the City*. MIT Press, Cambridge

Deogun J, Raghavan V, Sarkar A, Sever H (1997) Data mining: Trends in research and development. In: Lin T, Cercone N (eds), *Rough Sets and Data Mining – Analysis of Imprecise Data*. Springer, New York, 4–46. CrossRef.

Doel M, Hubard P (2002) Taking world cities literally. *City* 6[3]: 351–368. CrossRef.

Famili A, Shen WM, Weber R, Simoudis E (1997) Data preprocessing and intelligent data analysis. *Intelligent Data Analysis* 1[1]: 33–47. CrossRef.

Fayyad U, Piatetsky-Shapiro G, Smyth P (1996) From data mining to knowledge discovery: An overview. *Artificial Intelligence Magazine* 17: 37–54

Greco S, Matarazzo B, Slowinski R (1997) Rough set approach to multi-attribute choice and ranking problems. In: Fandel G, Gal T (eds), *Multiple Criteria Decision Making. Lecture Notes in Economics and Mathematical Systems, vol 448*. Springer, Berlin, 318–329. CrossRef.

Gregory D, Johnston R, Pratt G, Watts M, Whatmore S (2009) *Dictionary of Human Geography* (fifth ed.). Blackwell, Oxford

Kourtit K (2014) The new urban world, assessment of creativity, diversity and attractivity of cities. PhD dissertation, VU University Press, Amsterdam

Kourtit K (2015) The new urban world: Economic-geographical studies on the performance of urban systems. PhD dissertation, Poznan, Poland

Kourtit K, Nijkamp P (2013a) Creative buzz districts in smart cities: Urban retro-fitting and urban forward-fitting plans. *Romanian Journal of Regional Science* 7[2]: 37–57

Kourtit K, Nijkamp P (2013b) In praise of megacities in a global world. *Regional Science Policy and Practice* 5[2]: 167–182. CrossRef.

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

Kourtit K, Nijkamp P (2013d) The use of visual decision support tools in an interactive stakeholder analysis – Old ports as new magnets for creative urban development. *Sustainability* 5: 4379–4405. CrossRef.

Kourtit K, Nijkamp P, Franklin RS, Rodríguez-Pose A (2014) A blueprint for strategic urban research: The 'urban piazza'. *Town Planning Review* 85[1]: 97–126. CrossRef.

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

Kourtit K, Nijkamp P, van Leeuwen E (2013) New entrepreneurship in urban diasporas in our modern world. *Journal of Urban Management* 2[1]: 25–47

Mills ES, Hamilton B (1994) *Urban Economics* (fourth ed.). Harper Collins, New York

Mori Memorial Foundation (2012) Global power city index 2012. Tokyo: Institute for Urban Strategies at the Mori Memorial Foundation

Neal ZP (2012) *The Connected City – How Networks are Shaping the Modern Metropolis*. Routledge, New York

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

Orlowska E (2013) *Incomplete Information: Rough Set Analysis*. Physica Verlag, Heidelberg. CrossRef.

Pawlak Z (1991) *Rough Sets: Theoretical Aspects of Reasoning About Data*. Kluwer, Dordrecht

Pawlak Z (1992) Rough sets: Introduction. In: Slowinski R (ed), *Intelligent Decision Support – Handbook of Applications and Advances of the Rough Set Theory*. Kluwer, Dordrecht, 1–2

Pawlak Z, Grzymala-Busse J, Slowinski R, Ziarko W (1995) Rough sets. *Communications of the ACM* 38[11]: 89–95. CrossRef.

Pawlak Z, Skowron A (2007) Rudiments of rough sets. *Information Sciences* 177: 3–27. CrossRef.

Pawlak Z, Slowinski R (1996) Decision analysis using rough sets. *International Journal of Approximate Reasoning* 15[4]: 333– 365

Poel D (1998) Rough sets for database marketing. Physica-Verlag, Heidelberg, 324–335. CrossRef.

Polkowski L (2003) *Rough Sets: Mathematical Foundations*. Heidelberg. CrossRef.

Sassen S (1991) *The Global City*. Princeton University Press, Princeton. CrossRef.

Slowinski R (1995) Rough set processing of fuzzy information. In: Lin TY, Wildberger A (eds), *Soft Computing: Rough Sets, Fuzzy Logic, Neural Networks, Uncertainty Management, Knowledge Discovery*. Simulation Councils, San Diego, CA, 142–145

Slowinski R, Stefanowski J (1994) Handling various types of uncertainty in the rough set approach. In: Ziarko WP (ed), *Rough Sets, Fuzzy Sets and Knowledge Discovery*. Springer, London. CrossRef.

Taylor P (2004) *World City Network. A Global Urban Analysis*. Routledge, London

Tellier L (2007) *The World Urban History*. Presses Universitaires de Québec, Québec

Tranos E, Kourtit K, Nijkamp P (2014) Digital urban network connectivity: Global and Chinese Internet patterns. *Papers in Regional Science* 93: 409–428. CrossRef.

van den Bergh JCJM, Button K, Nijkamp P, Pepping G (1998) *Meta-analysis in Environmental Economics*. Kluwer, Dordrecht. CrossRef.

Vaz E (2016) The future of landscapes and habitats: The regional science contribution to the understanding of geographical space. *Habitat International* 51: 70–78

Vaz E, Arsanjani J (2015) Crowdsourced mapping of land use in urban dense environments: An assessment of toronto. *The Canadian Geographer/Le Géographe Canadien* 59[2]: 246–255

Vaz E, Cusimeno MD, Hernandez T (2015) Land use perception of self-reported health. *Land Use Policy* 46: 232–240

Vaz E, de Noronha Vaz T, Galindo PV, Nijkamp P (2014) Modelling innovation support systems for regional development–analysis of cluster structures in innovation in Portugal. *Entrepreneurship & Regional Development* 26[1-2]: 23–46

Wu C, Yue Y, Li M, Adjei O (2004) The rough set theory and applications. *Engineering Computations* 21[5]: 488–511. CrossRef.

Ziarco W (1998) Rough set as a methodology for data mining. In: Skowron A, Polkowski L (eds), *Rough Sets in Knowledge Discovering*. Physica Verlag, Heidelberg, 554–576

A Detailed Results

The frequency scores of the sub-sub factors in the six composite factors are recorded in Tables A.1 and A.2. When no number is shown, it means the variables are never part of a minimal reduct and therefore the least important ones in explaining the composite factors.

Table A.1: Frequency scores of the sub-sub factors of Economy

Economy	% frequency
GDP per Capita	100
World's Top 300 Companies	100
Office Area per Employee	100
Level of Economic Freedom	50
Number of Employees	50
Corporate Tax Rate	50
Nominal GDP	25
GDP Growth Rate	25
Total Market Value of Listed Shares on Stock Exchanges	
Number of Employees in Service Industry for Business Enterprises	
Wage Level	
Ease of Securing Human Resources	
Level of Political, Economic and Business Risk	

Table A.2: Frequency scores of the sub-sub factors of Livability

Livability	% frequency
Maturity of Community	100
Average House Rent	83
Medical Doctors per Population	83
Price Level	50
Healthy Life Expectancy Rate	50
Total Unemployment Rate	33
Population Density	33
Total Working Hours	17
Satisfaction of Employees for their Lives	17
International Schools per Foreign Population	17
Number of Murders per Population	
Variety of Retail Shops	
Variety of Restaurants	

Table A.3: Frequency scores of the sub-sub factors of Research and Development

Research and Development	% frequency	
	4 classes	3 classes
Number of Winners of Highly-Reputed Prizes	100	25
Interaction Opportunities Between Researchers	100	25
Number of Researchers	67	100
World's Top 200 Universities	67	100
Academic Performance in Mathematics and Science	67	25
Research and Development Expenditure	50	25
Number of Registered Industrial Property Rights (Patents)	33	
Readiness for Accepting Foreign Researchers	17	100

Table A.4: Frequency scores of the sub-sub factors of Cultural Interaction

Cultural Interaction	% frequency	
	4 classes	3 classes
Number of Theaters and Concert Halls	100	29
Number of Visitors from Abroad	100	57
Opportunities of Cultural, Historical and Traditional Interaction	100	
Number of Stadiums	50	14
Number of Large World-Class Cultural Events Held	50	57
Trade Value of Audiovisual and Related Services	50	
Number of International Conferences Held	25	43
Environment of Creative Activities	25	14
Number of Museums		100
Number of Guest Rooms of Luxury Hotels		100
Number of Hotels		
Level of Satisfaction for Shopping		86
Level of Satisfaction for Dining		100
Number of Foreign Residents		
Number of International Students		100
Number of World Heritage Sites (within 100km Area)		

Table A.5: Frequency scores of the sub-sub factors of Environment

Environment	% frequency	
	4 classes	3 classes
CO2 Emissions	100	100
Density of Suspended Particulate Matter (SPM)	100	100
Level of Green Coverage	100	100
Comfort Level of Temperature	100	100
Number of Companies with ISO 14001 Certification	50	100
Density of Sulfur Dioxide (SO2), Density of Nitrogen Dioxide (NO2)	50	100
Percentage of Renewable Energy Used		100
Percentage of Paper Recycled		
Water Quality		

Table A.6: Frequency scores of the sub-sub factors of Accessibility

Accessibility	% frequency	
	4 classes	3 classes
Punctuality of Public Transportation	100	
Density of Metro Stations	86	100
Cities with Direct International Freighter Flights	85	
Number of Cities with Direct International Flights	71	
Number of Runways	43	
Travel Time between Inner-city Areas and International Airports	43	100
Number of Passengers on International Flights	29	100
Convenience of Commuting	29	
Transportation Fatalities per Population		100
Taxi Fare		

REGION

The Journal of ERSA
Powered by WU

Volume 4, Number 1, 2017, 71–92
DOI: 10.18335/region.v4i1.172

journal homepage: region.ersa.org

Potentials and limitations for the use of accessibility measures for national transport policy goals in freight transport and logistics: Evidence from Västra Götaland County, Sweden

Anders Larsson[1], **Jerry Olsson**[2]

[1] University of Gothenburg, Gothenburg, Sweden (email: anders.larsson@geography.gu.se)
[2] University of Gothenburg, Gothenburg, Sweden and University of Oxford, Oxford, UK
(email: jerry.olsson@geography.gu.se)

Received: 28 October 2016/Accepted: 13 April 2017

Abstract. Swedish national transport policy treats freight transport as a major facilitator of economic development at all geographical levels. It is simultaneously noted that methods and data for business location and transportation are inadequate for following up transport policy objectives. This paper reports on a pilot study of the potential to develop accessibility measures to support and follow up policy objectives in the Swedish context. The accessibility concept and its practical application in concrete measures are discussed and problematized. Several practical examples from Västra Götaland County are used as illustrations. In terms of results, the study identifies several potentials and limitations of using accessibility measures to address freight transport issues. These measures' usefulness is driven mainly by the integration capability of transport and land use. This permits more complex questions and measures, supporting the integration between planning specializations. Limitations largely concern data availability and quality and the extent to which maps and measures can be communicated to non-experts. The concluding discussion highlights how the policy and governance context is central to understanding how best to utilize the potential strengths of the accessibility concept and related measures.

Key words: accessibility, freight transport, logistics facilities, transport policy

1 Introduction

Swedish national transport policy is intended to ensure that the transport system supports the country's economic development in a sustainable way, for example, by promoting the accessibility of logistics terminals, ports, and other transport infrastructure. It is simultaneously noted that methods and data for business location and transportation are inadequate for completely following up transport policy objectives in the field (Transport Analysis 2013a).

Shifts in the logistics sector have for quite some time illustrated a trend towards the consolidation of terminals, a trend expected to continue in the future (Hesse 2008, McKinnon 2009, Dablanc, Ross 2012, Sheffi 2012). This will lead to relatively fewer but larger units that may fundamentally change the accessibility of logistics services to businesses, especially those located outside major urban centres and transport corridors. It is therefore crucial that national and regional policymakers and planners better understand and follow up the shifting landscape of modern supply chains, as this underscores the importance of comprehensive efforts to coordinate land use, not only for logistics facilities but also other businesses that generate freight movements (Barysiene et al. 2015, David 2015, Sakai et al. 2015).

A key issue is the availability of data for both mapping current conditions and following change over time (Allen, Browne 2008). Freight flow data, for example, is only available at the crude geographical level of counties (Swedish Transport Administration 2013). One further explanation for the lack of comprehensive goals and targets is the heterogeneity of the freight transport market, in which each type of good has specific requirements and thus specific conditions and needs. This is sustained by the traditional institutional separation in which Swedish spatial planning is characterized by a two-level system based on the nation-state, at one level, and on local authorities with considerable autonomy, on the other. This system is not unique to Sweden (see David 2015 on land use planning in the United States). This has resulted in a division of spatial planning responsibilities, with strategic infrastructure and accessibility issues being handled at the national level, while local authorities are responsible for land-use planning, including business location. Although the last decade has witnessed growth in regional-level influence (Lindström 2007), there is still a substantial lack of coordination between actors at different geographical levels.

The Mobility Research Group at the Department of Economy and Society at the University of Gothenburg[1] has for several years worked on accessibility issues in personal transport, using a GIS-based model with high spatial resolution and connected to individual-level data (Elldér, Ernstson, Fransson, Larsson 2012). Several projects have been implemented, the most comprehensive one being implemented jointly with Västra Götaland Regional Planning Authority (Elldér, Gil Solá, Larsson 2012). This GIS-based model also has potential in the freight sector, as exemplified by recent research into city logistics and freight terminal utilization (Olsson, Woxenius 2012, 2014).

To pilot the methodological possibility of combining high spatial resolution with employment and industrial sector data, the Swedish government agency, Transport Analysis, has commissioned a project to test opportunities for better understanding regional patterns, accessibility potentials, and data requirements for accessibility analysis in freight transport (Olsson, Larsson 2016).

Within this context, the aim of this article is to *introduce and examine the potentials and limitations of accessibility analysis for monitoring policy goals in freight transport*. This scope can be specified in three research questions, as follows:

- What are the main strengths and weaknesses of the current accessibility approach relative to other similar concepts?
- How can accessibility measures be applied to freight transport problems?
- What are the most important strengths and limitations of accessibility analysis for monitoring transport policy goals?

The paper continues with a presentation and discussion of the accessibility concept and related measures, followed by a brief presentation of Swedish transport policy goals with a focus on freight transport. This presentation is then illustrated by examples of measures applied to Västra Götaland County. The results of the mapping and database analysis are brought together in the discussion section, which is followed by the presentation of conclusions.

[1] http://es.handels.gu.se/english/units/unit-for-human-geography/research/research-groups/mobility-research-group

2 From mobility to accessibility planning: concepts and measures

2.1 Introduction

Recent decades have witnessed increasing travel and increasing levels of car dependence in cities worldwide (Handy 2002, Banister 2005, 2008). Mobility has become taken for granted in people's everyday lives and in business strategy development. This has profound implications for the flows and relationships that build cities and regions. Simultaneously, cities are experiencing a range of mobility-related problems, such as congestion, lack of safety, noise, and air pollution, all imposing societal costs and affecting business productivity (Bertolini 2005, Melo et al. 2009).

These are obviously problems whose solutions call for an integrated, holistic approach. Traditional planning perspectives often separate the fields of infrastructure, land use, and business development, treating them as "silos" with little or no coordination. The concept of accessibility brings together transport/mobility and land use, offering a useful framework for more integrated planning strategies (Geurs, van Wee 2004, Ferreira et al. 2012). It also offers a potentially powerful guide that planning practitioners can employ to develop and test effective strategies for building sustainable cities and regions (Straatemeier 2008).

2.2 The accessibility concept

In the following, we refer to accessibility based on the following definition: "The extent to which the land use–transport system enables (groups of) individuals or goods to reach activities or destinations by means of a (combination of) transport mode(s)" (Geurs, Ritsema van Eck 2001, p. 19). This definition is based on four components:

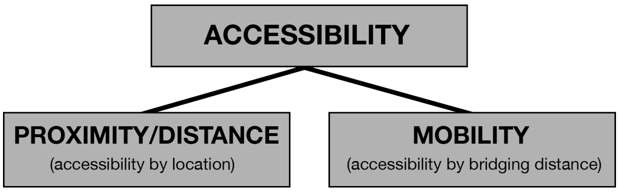
- *Transportation*: the cost (e.g., time or monetary costs) of moving goods or people physically to a destination including all modes of transport.
- *Land use*: the location and geographical distribution of potential origin and destination points, such as population density, jobs, services, customers, and warehouses.
- *Time*: access to facilities and activities varies during the day depending on, for example, the operating hours of suppliers and customers.
- *Individual*: individual values, opportunities, needs, and constraints represent important conditions for the broad application of the accessibility concept.

In an early application to spatial planning, Hansen (1959) broadly defined accessibility as: "the potential of opportunities of interaction". This suggests two contrasting definitions of accessibility, one addressing how points in a network relate to each other (i.e., access to potential opportunities from one specific starting point), and another depicting how all points are related to all other points on a surface (i.e., general spatial potential). Ingram (1971) later referred to this distinction as being between relative and integral accessibility, respectively. Our point of departure is the relative approach, since we are interested in accessibility for specific purposes, travelling to/from specific activity points. Before exploring accessibility measures in greater detail, we would like to draw attention to several related concepts.

2.3 The relationship between accessibility, mobility, and proximity

Accessibility as defined above is related to the concept of mobility, since movement is a means to achieve accessibility. It includes all forms of physical movement using or not using vehicles as well as virtual movement using communication technology. Historically, we have witnessed progression towards increasing car-based mobility over gradually longer distances in order to facilitate accessibility. However, given current discussion of mobility and sustainability (Banister 2008), it is important to highlight the opposite strategy, namely, accessibility through proximity.

Figure 1 illustrates in a simplified way two principal approaches to achieving good accessibility, i.e., spatial proximity and overcoming distance. Achieving good accessibility



Source: Haugen 2012

Figure 1: The interrelationship between accessibility, mobility, and proximity

through proximity emphasizes the land-use component, with the importance of the location of points of production and consumption. Any improvement of accessibility then involves relocating either the origin or destination points, or both. The traditional policy solution to poor accessibility has been to concentrate on mobility by improving road infrastructure capacity, creating the potential to uphold reliable but potentially unsustainable transport systems over long distances.

This highlights a potential conflict between traditional transport-based planning and current land-use-based ideas of proximity- and non-car-based planning in cities (Newman, Kenworthy 2015). This debate has emanated from person-transport considerations but is as important for freight. We argue that this is one area where an accessibility approach provides a platform for the integration of transport and land use. It is important to note that normative measures are often needed in practical planning contexts (Páez et al. 2012), possibly leading to a situation in which opposing policy goals for transport and land use are present in the same accessibility measure.

These concepts and distinctions are fundamental, not only in an abstract theoretical sense but also in order to understand everyday planning problems. One example from personal transport is the so-called accessibility paradox (Haugen 2012, Haugen et al. 2012, Haugen, Vilhelmson 2013), in which measured reduced average travel times between residents and the nearest locations of important daily activities (e.g., schools and services) do not always correspond to shorter actual trip distances. Other factors and preferences (e.g., type of school and education) can be seen as more important than proximity. This illustrates that it is far from obvious that the relationship between accessibility, mobility, and distance is determined by a one-dimensional rational choice to minimize either distance or time. This is even more problematic in freight transport, in which decisions about route choice or terminal usage are made by several layers of actors and very rarely based on a shortest-distance logic (Melo et al. 2009).

Given this discussion, planning for improved accessibility in freight transport is complex. For example, it is important to critically examine for whom and where accessibility will be improved. Different firms at the same location might have very different accessibility needs and interests due to, for example, their types of goods, business sector, logistics management solution, customer density, and traffic congestion situation and to the effect of freight transport on local residents (Melo et al. 2009).

2.4 Overview of accessibility measures

To move from accessibility principles to implementation, we must look more closely at the most common accessibility measures. Geurs, van Wee (2004) identified four types of accessibility measures depending on the perspective or problem area considered: i) infrastructure-based, ii) activity-based (divided into location and space–time measures), iii) individual-based, and iv) utility-based measures. It is important to recall that, depending on the approach chosen, it is only possible to address a limited set of problems. Since this paper applies mainly location-based measures, we will emphasize these. However, a brief overview of other types of measures is given first.

An infrastructure approach to accessibility focuses on the properties of road, rail, bike, and walking networks, such as their connectivity and speed. These network properties provide a good indication of accessibility and mobility near where roads or railways and their nodes are located, but say less about accessibility over a wider hinterland. It is also difficult to measure, for example, the effect of increased road speed, as this is directly related to land-use factors such as population, jobs, and commuting (Geurs, Ritsema van Eck 2001).

Individual-based measures start at the individual level and consider the opportunities and limitations in performing activities at different locations during a given period, such as one day. The approach is largely based on a time-geographic approach (Hägerstrand 1970) in which the abilities of individuals or groups (e.g., households) to perform daily activities and projects are determined by individual and joint restrictions in time and space. Accessibility then becomes more complex than just being able to reach a certain point within a certain optimal distance or time. It also involves relationships with all other events and locations that need to be temporal-spatially coordinated during the day. When applied to freight transport and logistics, this approach might be used to better address limitations such as the opening hours of firms/suppliers/terminals and peak-hour congestion. This approach might be more applicable in a business context than in public land-use planning.

Based on economic utility theory, one can use businesses' valuations of the benefits of various alternative destinations as an accessibility measure. The assumed logic here is that decision makers make rational choices intended to maximize their own cost/benefit ratio. The advantage of this approach is that the measure has a strong theoretical coupling to economic theory. This is also its drawback, as considerable theoretical knowledge is required in order to understand such an abstract measure and how it can be implemented and communicated in practical planning (see La Paix, Geurs 2016).

2.5 Location-based measures

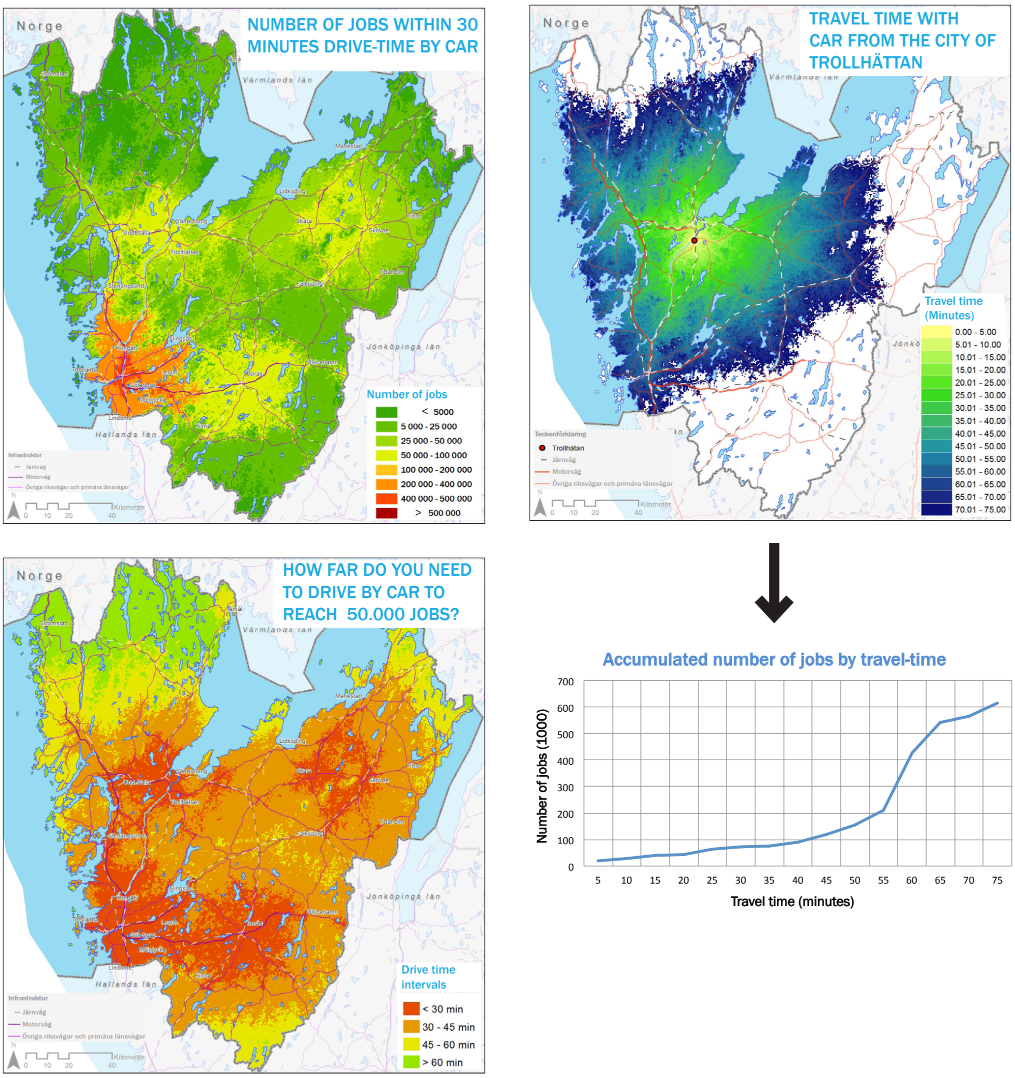
Location-based accessibility measures take into account both the transportation and the land-use components. In local or regional planning, this type of measure is considerably more relevant and is more widely applicable since it is directly linked to particular geographical areas and their demographic and socio-economic contexts. Such measures can range from straight-line distance or travel-time measures from one starting point to one or more destinations, to complex implementations of distance friction and the weighting of destination importance and competition.

The most straightforward measure is distance, either in absolute terms, such as the number of metres, or in relative terms, such as travel time. The aim is often to calculate either the number of potential customers or suppliers within a given catchment area or the closest destination(s) of interest. If we instead want to know how many warehouses (or other destinations) can be reached within a specific time interval, cumulative measures (i.e., contour or isochrone measures) are used. Breheny (1978) divided the accessibility of opportunities measure into three dimensions: starting points, destinations, and costs. By keeping one factor constant, it is possible to obtain a relatively easy-to-understand cumulative measure:

- *Keeping costs constant* measures the opportunities available within a certain cost in distance, time, or monetary terms; for example, the number of potential jobs reachable within a 30-minute trip.
- *Keeping target points constant* measures the cost of reaching a predetermined number of opportunities; for example, the average travel time across a region to reach 50% of the warehouses in that region.
- *Keeping the starting point constant* measures the potential opportunities based on variations in cost; for example, the number of warehouses within different travel-time intervals from a selected point.

These types of measures give results in the form of spatial distributions, allowing visualization in the form of influence fields or hinterlands. Figure 2 illustrates the spatial outcomes of these measures.

These metrics are useful in practical planning due to their relative simplicity and resulting communicability to a wider audience. This simplicity also presents problems,



Source: sources: authors' calculations; road network, Swedish Transport Administration; other map data, Swedish Land Survey; socio-demographic data, Statistics Sweden

Figure 2: Illustrations of different accessibility measures using the same data: fixed cost (top left), constant target points (bottom left), constant starting point (top right), and potential jobs by travel time (bottom right)

however. An important limitation is that the used measures treat all target points as equally valuable, regardless of the travel time. It is important that the travel time and distance parameters be determined so as to reflect actual travel behaviour. For freight transport, this is a major limitation, since there is no direct equivalent to travel surveys in personal transport studies.

The value of distance can more realistically be calculated using functions to capture friction of distance, meaning that the farther away something is, the less its influence on the final result. The advantage of such metrics is that distance affects accessibility, enabling calculations based on empirical studies of actual travel behaviour. This yields more accurate and theoretically consistent results, which, however, may be difficult for outsiders to interpret. This approach also requires that function parameters be based on empirical travel surveys for them to be reliable for a specific region.

3 Swedish national transport policy goals for accessibility and freight transport

The overall Swedish national transport policy objective states that the transport system should contribute to economic development throughout the country. In terms of accessibility, this requires good access to logistics terminals, ports, and other road/rail transport infrastructure. At the same time, concerning the monitoring of transport policy objectives and methods, it has been found that access to data on business location and transportation is inadequate (Transport Analysis 2013b,c).

One basic prerequisite for the creation of measures that contribute to the more efficient use of transport system resources is reliable information on the location of freight transport-intensive operation facilities. Authorities with overall responsibility for the follow-up of policy goals, such as Transport Analysis (2013b,c), need quantitative information on the accessibility of important logistical destinations, such as ports and logistics centres, to various industries. There is currently a lack of appropriate indicators, as the disaggregated data are of insufficient quality.

A question, then, is whose perspective should be taken into account? As logistics and freight transport comprise disparate activities with varied requirements and conditions, different actors' logic, behaviour, and priorities frequently counteract each other. There is a risk of sub-optimization in which all actors streamline their transportation needs based on the conditions governing their specific operations. When these actors are highly autonomous, individual decisions are made without paying attention to the entire logistics chain (Ministry of Enterprise and Innovation 2011).

The Energy Commission's final report states that long-term investments in infrastructure and other planning need to target the creation of conditions for a more energy-efficient transport system (SOU 2008). From a national perspective, therefore, knowledge of business locations and of the accessibility of key nodes and hubs (including ports and airports) is central, as both long-term planning and investment are expected to concentrate on fewer transport nodes and designated freight transport corridors, as illustrated in the following quotation (Swedish Transport Administration 2013):

> In the national plan for the transport system, a strategic network for long-distance haulage is defined. The network includes a designated road/rail network, central intermodal terminals, ports, and airports. This network is considered to be of paramount importance for long-distance business freight transport. Therefore, Transport Administration actions should be concentrated on the strategic network. (Swedish Transport Administration 2013, p. 12)

This need to define a national core network follows partly because the distances within the country are great and partly because Sweden, as a major exporter, is very sensitive to the efficiency of increasingly spatially extended global transport networks. From a national perspective, industrial logistics costs are therefore central, from both the short-term competitive and long-term sustainability perspectives. However, the societal perspective has much broader significance, and the overall transport policy objective is to ensure economic efficiency for both business and citizens throughout the country. According to the Ministry of Enterprise and Innovation (2011), a priority is to increase the efficiency of freight transport and optimally use the existing infrastructure (increased infrastructure capacity cannot itself meet the entire expected increase in transport demand). From a societal perspective, therefore, logistics must be recognized as transportation activity wherein capacity, traffic environment (i.e., safety), accessibility, and energy consumption are taken into account for all users – including passengers. At the same time, these considerations need to be balanced by a good economic and business climate. Taken together, this calls for clear indicators for monitoring community investments in transportation and land use.

4 Accessibility measures and their application in freight transport: examples from Västra Götaland County

4.1 Method, material, and limitations

Accessibility measures and maps are calculated using purpose-built software initially developed for personal transport to combine advanced transport modelling of public transport with GIS functionality (Berglund 2001). This accessibility tool combines georeferenced data with information about the transport system and travel options. It enables the computation of accessibility by both car and public transport of many potential supply points, such as amenities and workplaces.

For this study we have used only the capabilities related to movement along the road network. The computations were based on calculated travel times between 500×500-metre cells, allowing analysis beyond administrative divisions. For a more in-depth description of the tool functionality see Elldér, Ernstson, Fransson, Larsson (2012) and Larsson et al. (2014).

Each cell contained information on the number of logistics terminals and food-sector firms. Logistics terminals are defined as facilities for the reception, delivery, consolidation, distribution, and storage of freight in the transportation process. Food-sector firms comprise facilities covered by the Swedish industry classification code SNI 10, which covers the production of foodstuffs. This allows us to link driving times, land uses (or location patterns), and economic indicators on the production side of the logistics-sensitive food sector. We are aware that the distribution side of the food sector is also highly relevant in terms of logistics restructuring. However, a lack of detailed data on retailers and large-chain internal terminals limits the potential to include such information in this illustrative case. This issue is elaborated on in sections 5 and 6.

The data used in the analysis is drawn from several sources, as follows:

1. *Road network* data, including speed limits for calculating driving time, was supplied by the Swedish Transport Administration's National Road Database (NVDB) and captures the situation in September 2014.

2. Data on *logistics terminals*, including their exact locations, was updated by the authors based on a previous study by the government agency Traffic Analysis (WSP 2013). The services and activities that terminals offer can be captured only partially using statistical industry codes. These terminals may range from storage-only facilities to complete logistics centres incorporating a wide range of complementary services. Substantial time was spent manually qualifying the function of terminals.

3. Data on *multimodal rail–road terminals*, including location and capacity, was collected by the authors from company and public authority websites.

4. *Locations of food-sector firms* were extracted from the internal GILDA longitudinal database comprising official register data provided by Statistics Sweden. Individuals are linked to their workplaces, making it possible to identify the type of activity associated with all workplaces in the region.

The measures are calculated based on certain conditions and limitations due to incomplete data or limitations in the accessibility model. In this study, logistics terminals are limited to those open to all potential commercial customers. This definition is applied because it reflects the true geographical potential. We can find many additional facilities within company-internal logistics systems. However, these are not potential service points for outsiders even if located within reach according to accessibility calculations.

In some analyses we refer to large terminals, which are terminals with annual turnovers of SEK 100 million in 2014 (equivalent to USD 12.5 million). This demarcation is relevant in order to identify sites with enough capacity to supply customers with a variety of services.

As discussed elsewhere in this paper, cumulative accessibility measures are dependent on the setting of time limitations. In this case, 60 minutes of driving time is used as the threshold in several maps. This is based on the principle of centrality, whereby a terminal's hinterland comprises roughly 100 km (Rodrigue et al. 2013, p. 135). Exactly where the line is established on a commercial basis may vary between carriers.

A further distinction concerns potential customers. Different business sectors have very different requirements in terms of distances, terminal facilities, etc. In this study, we use the food industry as a proxy for logistics-intensive industries, representing the part of the economy that is dependent on flexible and frequent logistics chains for daily activities and that could potentially use the terminals. One important limitation of the calculations is that they include only terminals and food-sector production plants in Västra Götaland County. This limitation was made due to the very time-consuming process of identifying logistics terminals, elaborated on in section 5. Since both logistics providers and terminals compete, this limitation creates uncertainty as to the results in border regions. This is especially significant in the eastern part of the region, where a significant market is reachable by 30 minutes of driving.

We would also like to raise the issue of data categorization on maps. Visual perceptions, and hence the analytical values, of datasets are highly dependent on the selection of appropriate classes and categories (Evans 1977). For this paper, we selected classes according to the principle of simplicity in communication with non-expert groups; for example, travel times are divided into 15-minute intervals. It is important to note that other classification schemes might reveal a more extreme hierarchical pattern in the data (Jiang 2013), being more suitable for research questions comparing highly accessible areas with less accessible ones.

4.2 Accessibility measures based on driving time to nearest facility

The most straightforward geographical measure is driving time to the nearest terminal. Figure 3 shows the driving time to the nearest terminal, regardless of its size or function. The map clearly illustrates how virtually every part of the region has access to a terminal within 30 minutes.

The map in Figure 4 illustrates the calculated accessibility to large terminals expected to provide the range of services needed for customers not part of corporate logistics networks, such as those of Volvo or IKEA. Using only the major terminals as destinations, differences in driving time become more pronounced. Moreover, it is possible to discern several corridors along major highways with short travel times. This also illustrates the usefulness of disaggregated spatial units: the use of more aggregated spatial units, such as municipalities, could cause the corridor effects shown in the map to be overlooked.

Driving time is useful due to its uncomplicated form. However, to benefit from accessibility analysis, there is a need to add the land-use component. This is illustrated in Figure 5 by the number of companies in the food industry located within different driving-time intervals.

Adding the potential number of users in the logistics-sensitive food sector illustrates the concentration and co-location of warehouses and food-sector firms. Half of the firms are under 15 minutes from the nearest terminal and more than 80 % can be reached within 60 minutes. The difference between large terminals and all terminals widens with increased travel time, but does not exceed 10%.

The use of driving time as a measure has several advantages, mainly because the result in minutes is an absolute value and very easy to communicate. This indicator is directly related to everyday freight transport issues such as delivery time and precision. A further advantage is that it is relatively easy to calculate, needing only two datasets: road network and terminal locations. This makes it affordable to calculate and, from a data supply perspective, it is relatively easy to collect longitudinal datasets in order to monitor changes over time.

There are also limitations to using driving time as a measure. Its simplicity means that it addresses only a very specific question. For following up the Swedish national transport policy objectives, however, this measure is insufficient on its own. A further drawback is the use of travel time without weighting. For personal transport measures, it is common to add weights to the points of attraction and to use distance decay functions for the travelling. Here the logic is that the larger the city, the greater the attraction, and conversely, the farther away, the greater the resistance. This procedure is supposed to

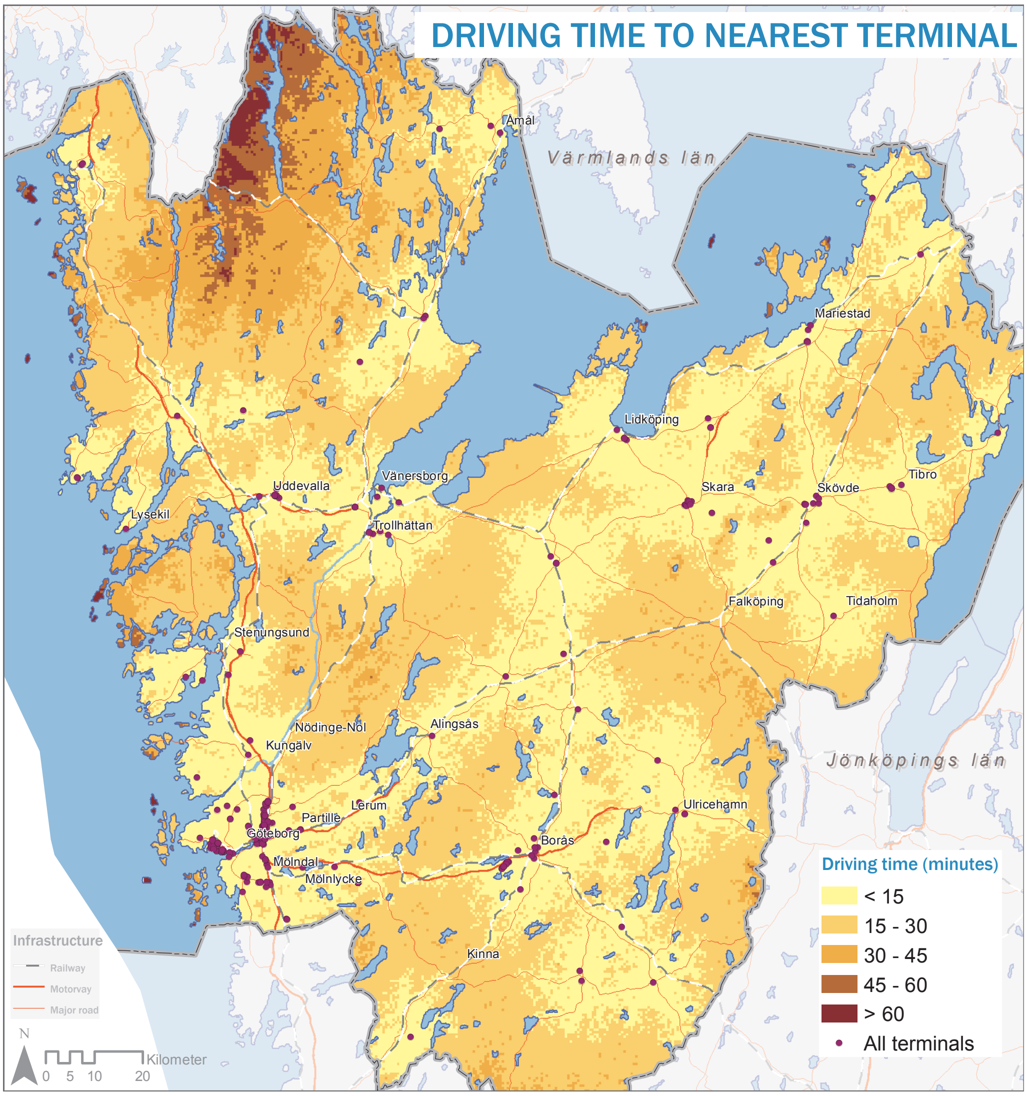


Figure 3: Accessibility in terms of driving time to the nearest terminal

make the analysis more realistic in relation to travel behaviour. The problem in freight transport calculations is lack of knowledge (mainly due to lack of primary data) of how goods actually move and of the criteria for the choices made. More knowledge of the "travel behaviour" of freight would be needed in order to determine whether, for example, the nearest terminal is actually a valid measure in real everyday situations.

4.3 Measures based on potential accessibility and catchment areas

While the measure used in Figures 3 and 4 illustrates accessibility from a transport buyers' perspective, the logic can be reversed and instead used to investigate the potential catchment areas of terminals, measured as the number of destinations reachable within a specific time. In this case, we have chosen 60 minutes of driving time.

The main advantage of this measure is that it connects infrastructure and space via potential, providing a clear basis for assessing what areas meet a certain objective or minimum requirement level. This measure is also relatively easy to communicate. Compared with the driving time indicator, catchment areas permit us to follow up municipalities' and regions' competitive potential in terms of how well they are situated in the terminal landscape.

However, one important limitation is the fact that the measure considers all destinations

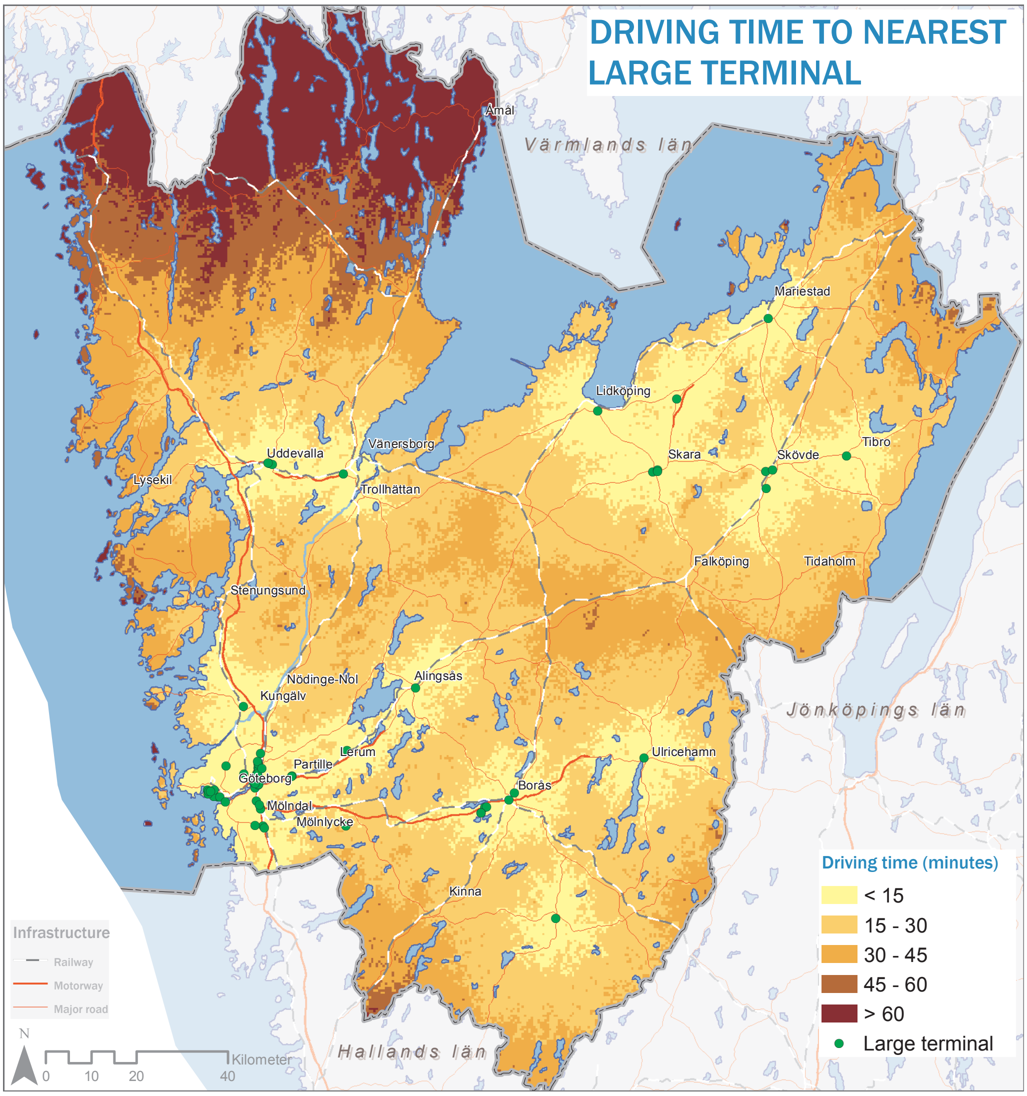


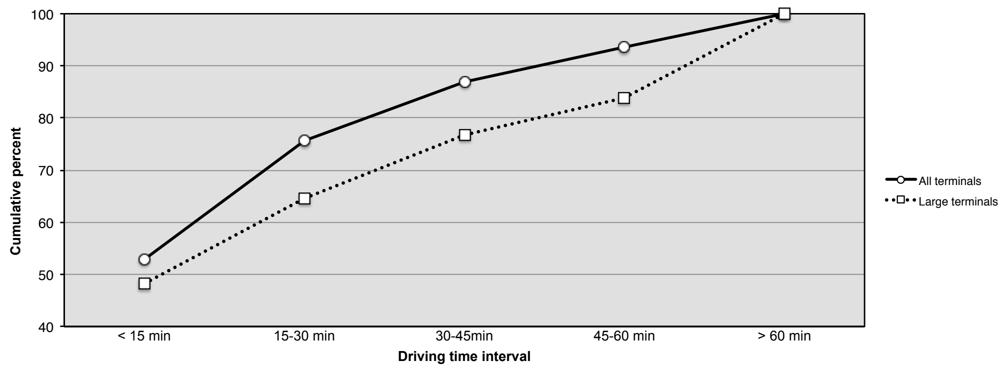
Figure 4: Accessibility in terms of driving time to the nearest large terminal, defined as having annual turnover exceeding SEK 100 million

equally valuable regardless of driving time. A terminal located a 60-minute drive away from the starting point is given the same value as a terminal located 5 minutes from the customer. Again, it is important to set driving time and distance parameters based on empirical data, to reflect real behaviour as well as possible.

Transferring the concept of travel behaviour to freight transport is complex, since the choice of route or terminal is usually not determined by the driver, but instead results from the commercial need to transport a product. The choice of starting point (e.g., firm) and destination point (e.g., terminal) then depends on several factors, such as business arrangements limiting what terminals to use and goods requiring special treatment (e.g., frozen foods). This brings the discussion to the question of constructing weights for destinations and travel distance to model accessibility more realistically.

4.4 Weighed measures

Some form of weighting is traditionally used in transport models (de Dios Ortúzar, Willumsen 2011) to create better theoretical approximations of actual travel behaviour. Destinations located farther away then become more costly to reach, while the more important the destination is, the more attractive it becomes. Translated into the terms



Source: authors' calculations; firm data from GILDA

Figure 5: Cumulative share of food industry companies within different driving-time intervals from all terminals and large terminals

of Figure 3, accessibility would not be determined solely by the travel time in minutes, but the minutes would become "longer" the farther one travels due to distance decay. In addition, the importance of each terminal in terms of, for example, size and turnover, would further influence the degree of attractiveness.

However, this approach has at least two important limitations. As noted above, we currently know very little about the "travel behaviour" of freight and therefore have limited empirical data with which to weight both distance decay and the relative influence of destinations. What we know is that transportation to and from terminals is part of increasingly complex supply chains. This means that the choice of terminal for a single shipment is determined by decisions much earlier in the process than when the transport is performed. A second limitation is the usefulness of the calculation results. Since both the distance and target point affect travel time based on various often nonlinear functions, the end result cannot be interpreted in minutes; instead, there is a need to construct relative weighted accessibility measures.

Figure 7 illustrates the driving time to the nearest terminal weighted by the terminals' annual turnovers. The weights are set using a simple linear logic in order to clearly illustrate the effects of a weighted measure. Small terminals with large hinterlands clearly illustrate poor relative accessibility. These results indicate that choosing the nearest terminal is often unrealistic, as exemplified by the two terminals located in Götene (see inset map in Figure 7). The southern terminal has a high turnover and is therefore favourably weighted, while the northern terminal has a considerably lower weight. In practice, there is only 1 minute of driving time between the two terminals, so all potential customers in the northern area of very extensive "weighted travel time" can easily choose the southern option instead. One method to solve the problems associated with treating the nearest terminal as the default would be to calculate the travel time to the nearest two or three terminals. A problem with this approach would arise in areas with very few terminals where the second or third terminal could be located very far away.

It is also possible to use weighting in analysing potentials at different locations. Figure 8 illustrates this with a calculation of accessibility weighted by terminal turnover, where higher turnover implies a higher potential service level. The map shows the aggregated total terminal turnover accessible within 60 minutes of driving time. The actual aggregated values for highly accessible areas become astronomical, obviously with no connection to the real values of points in the region. The advantage of this measure is that it qualifies the differences between areas with many terminals in absolute terms, but differing greatly in size. The pattern in Figure 8 does not differ significantly from that in Figure 6, but the values are much more dispersed and thus highlight the significance of being nearer large facilities, rather than just any facility.

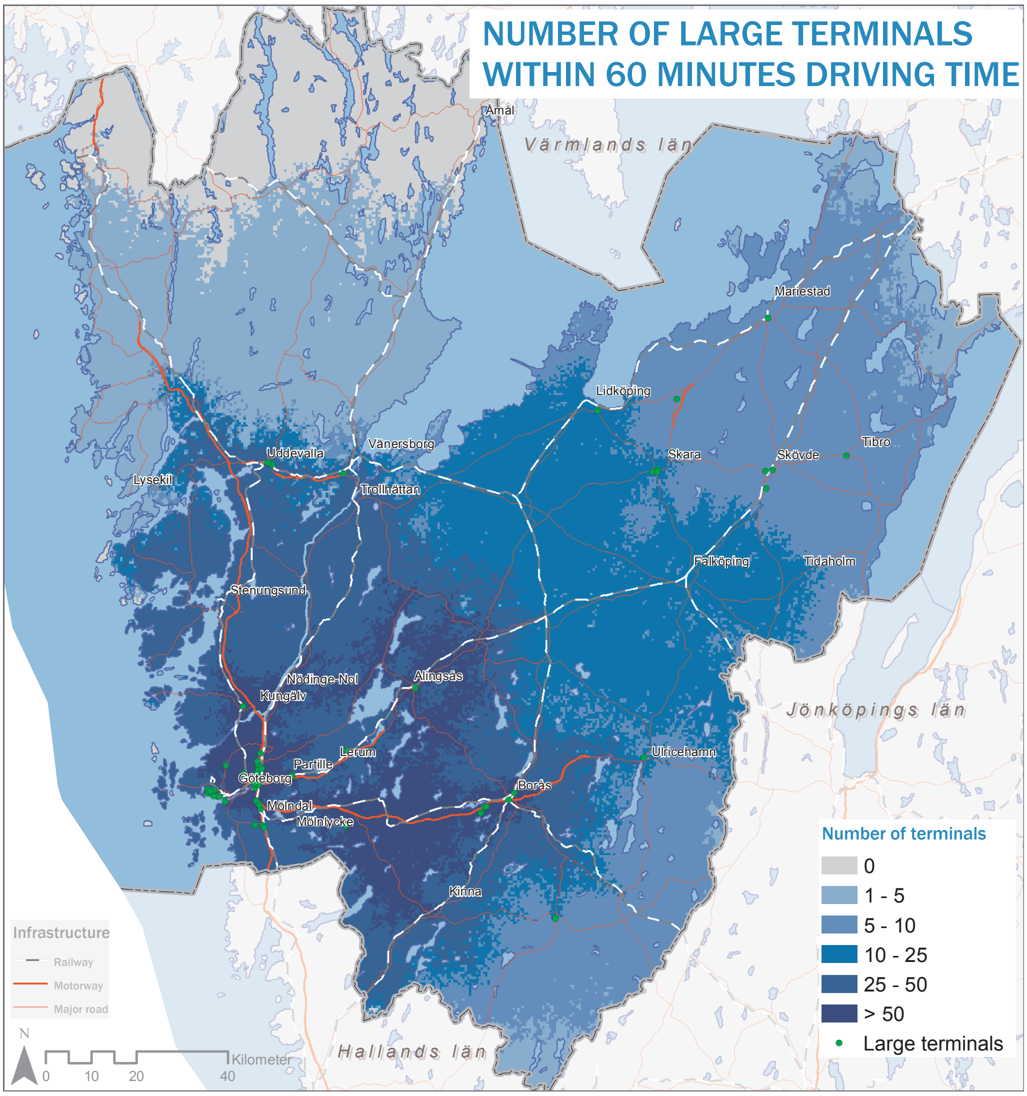


Figure 6: Number of large terminals accessible within 60 minutes of driving time

4.5 *Spatial aggregation of detailed measures*

Analysis and measures have so far been visualized with a spatial resolution of 500 metres. This resolution shows intra-regional variations very well, but is less suitable for follow-up at the national level. Since national targets are already using municipalities as the spatial unit for monitoring transport policy objectives, we propose using a method of spatial aggregation and combining measures into one single composite measure. The following exercise should be regarded as an example for discussion, rather than as a complete and finished measure ready for application.

The starting point for the measure will be the municipality. The first step is to assign a centre point to each municipality where trips start or finish depending on the measure. We suggest that centre points should be calculated based on the spatial pattern of logistics-sensitive industry. In this case, the weighted centre point for the food industry is our proxy for logistics-intensive industry.

We then select three measures meant to complement each other:

1. *Terminal proximity*: driving time from the municipal centre to the nearest terminal weighted by the terminal's economic turnover – This measures the extent to which the municipality has large terminal capacity in its immediate vicinity.

2. *Terminal potential*: the sum of the terminal's financial turnover earned within 60

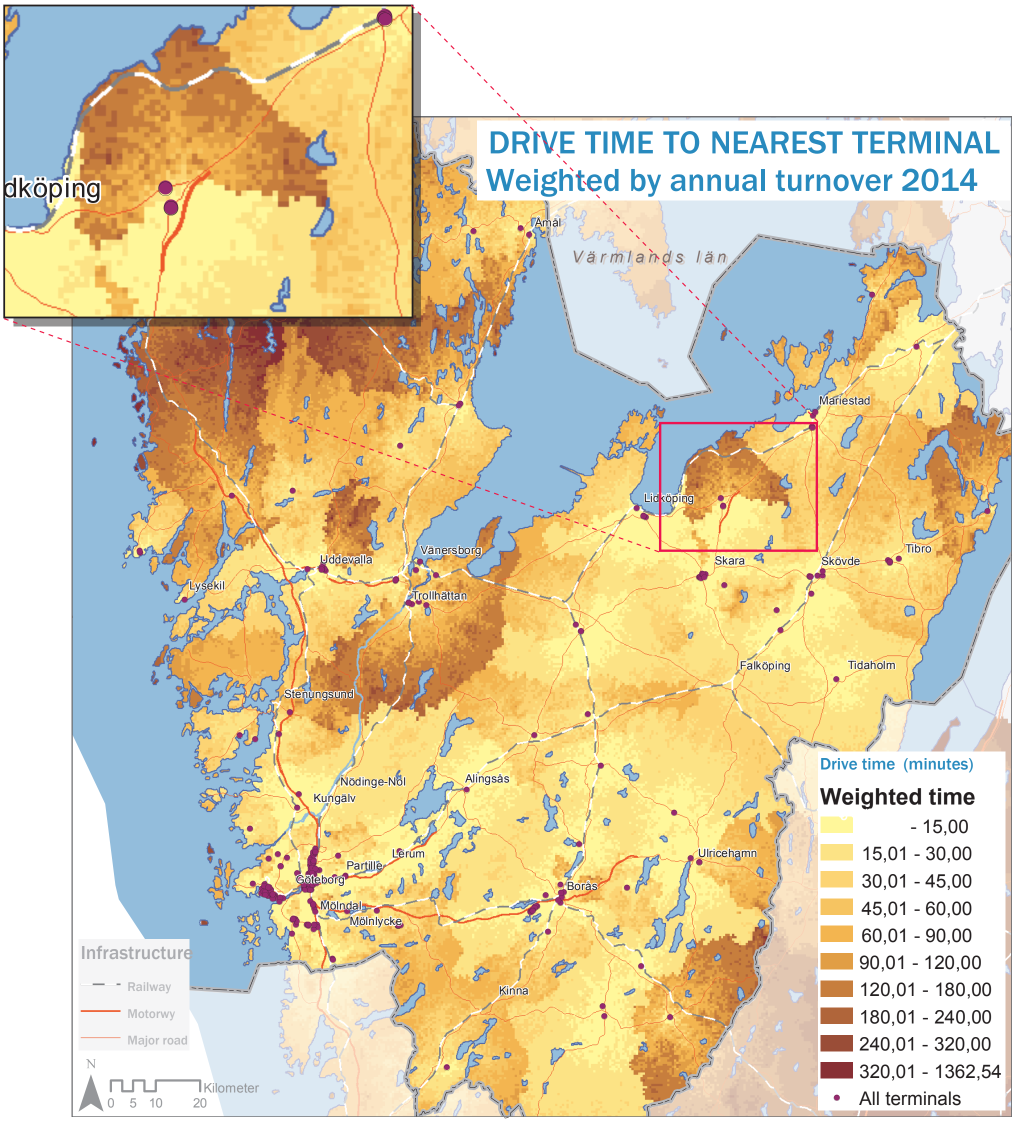


Figure 7: Accessibility measured as driving time to nearest terminal, weighted by annual turnover of terminal

minutes of driving time from the municipal centre – This complements terminal proximity and indicates the potential attractiveness of the area for the localization of logistics-intensive businesses.

3. *Road/Rail combination potential*: the total weighted value of the size of intermodal terminals reachable within 60 minutes of driving time from the municipal centre – This indicates the potential accessibility of good combined transport opportunities.

Each indicator is mapped in Figure 9. Indicator 1, the proximity indicator, produces more heterogeneous results than do the cumulative potential indicators illustrated in the two other maps. Accessibility is very good in the Gothenburg metropolitan region and along the major transport corridors. Also apparent are single peripheral municipalities each with a large terminal.

Indicator 2 reflects the dominance of the Gothenburg region due to the concentration of large terminals there. Again, note that the maps include only terminals located in Västra Götaland County, putting outlying municipalities at a relative disadvantage. Indicator 3 clearly shows the importance of Gothenburg; this particularly applies to the large multimodal terminals in the port, which weigh heavily in this context.

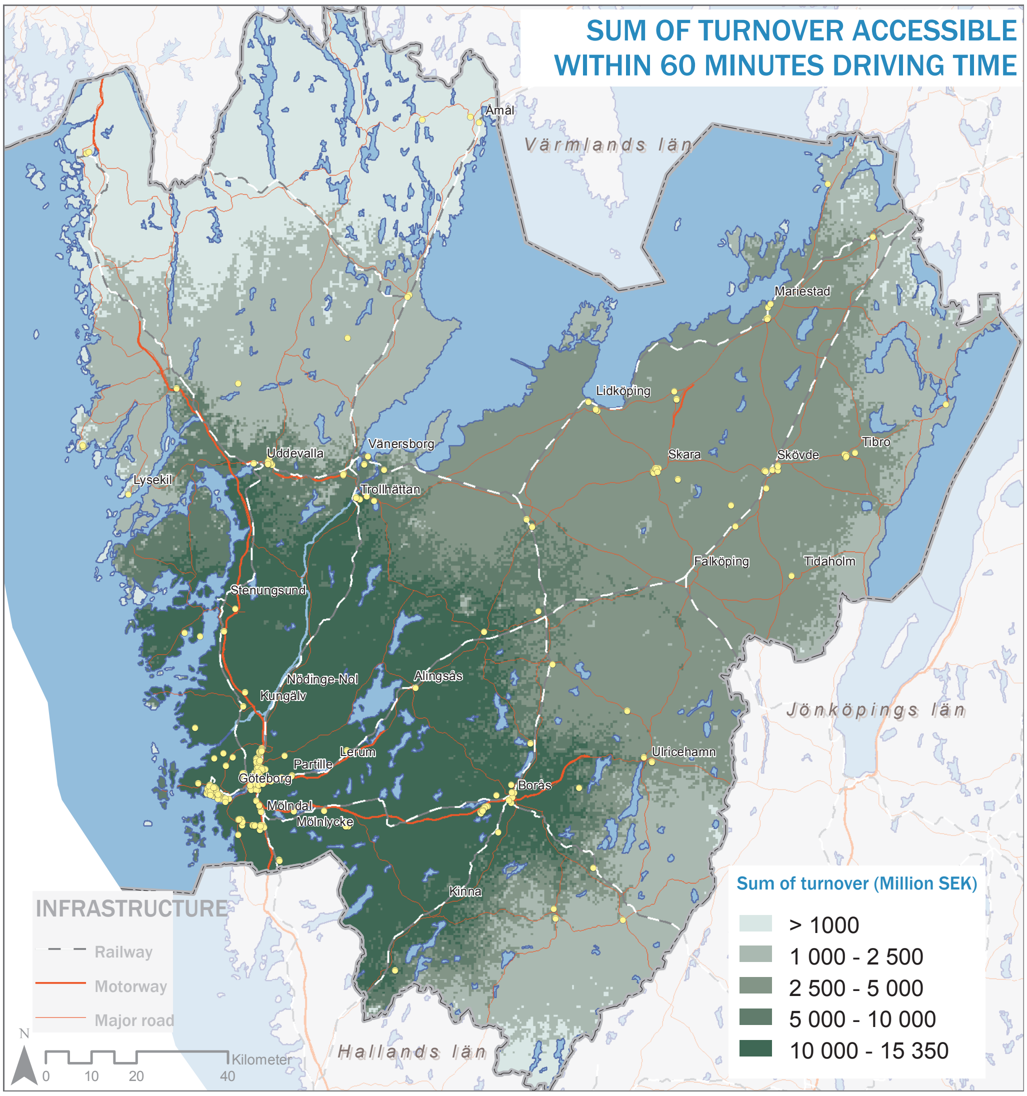


Figure 8: Aggregated terminal turnover accessible within 60 minutes of driving time

The measures are relatively straightforward in that they are related to driving time measured in minutes. However, merging the three indicators into one is problematic since they are weighted differently and their numerical outputs differ greatly. One way to achieve this would be to calculate the share of each municipality based on the total of the region and then add the relative numbers. Another method would be to rank all the municipalities for each indicator and then merge the three ranking figures to form a common ranking score. This method will be used here based on the simplicity of communicating and follow-up over time.

Figure 10 shows the summary score per municipality, illustrating how the major transportation corridors in the region have better access to logistics services than do their neighbouring municipalities. It is also very clear how proximity to the county's main urban centre of Gothenburg exerts an influence. One can see an inner and an outer ring of municipalities with very good accessibility located around Gothenburg.

5 Discussion

The main advantage of the accessibility concept in freight transport planning and policy lies in its integration of transportation and land use. Current approaches often focus on just one of the two. Transport planning is mostly concerned with network-related aspects

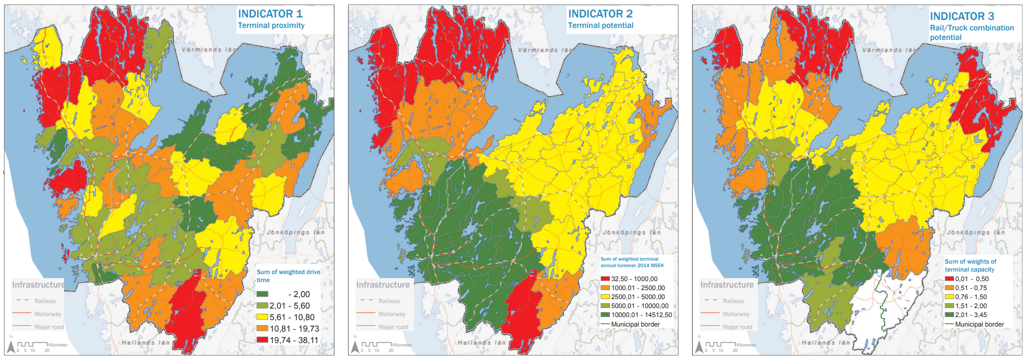


Figure 9: Illustration of the three municipality-based indicators. Left to right: terminal proximity, terminal potential, and rail/truck combination potential

of mobility efficiency such as speed, connectivity, flows, congestion, and parking. Land-use planning, in contrast, traditionally addresses issues related to human activities and their spatial outcomes, such as housing, economic development, and leisure activities. The accessibility concept offers a way to link these fields, taking a step towards a more integrated planning process (Bertolini 2005). It supports the link between transport/mobility and social sustainability issues (Lucas 2012), since changes in mobility always have different spatial accessibility outcomes for different groups in society. Being able to combine two areas of planning into one also entails potential problems. Obviously, an extra degree of complexity results when dealing with transport and land use simultaneously. Measures of accessibility will undoubtedly be more multi-dimensional than, for example, population density or traffic flows. There is also an added component when evaluating accessibility change over time, since both transportation and land use can change simultaneously over the same period.

Given the potential for the accessibility approach to support integration in planning, how can this approach be applied specifically to freight transport? This is an important question since freight and personal transport differ substantially in this respect. Unlike personal transport, freight transport is driven by derived commercial demand emanating from the business sector. Choice of transport mode or route is the combined result of a logistics chain comprising several frequently conflicting dimensions, such as cost, reliability, and scale economies. Accessibility in freight transport is much more likely to be a business matter, while in personal transport it directly mirrors the potentials of individuals and their everyday lives in different locations. Having this distinction in mind is important when asking questions about accessibility levels, for whom and where. On the other hand, accessibility measures may also be used with specific planning questions in mind, to understand how different groups in different places are, for example, influenced by new investment in logistics terminals and infrastructure.

Even a basic travel-time map provides a wealth of information. The use of “real time” measured in minutes without weighting or model calculations constitutes a potentially powerful outcome due to ease of interpretation: no expert knowledge is needed to conclude that ten minutes is twice as long as five minutes. This makes simple accessibility measures useful tools in communication between professional groups, such as planners and logistics business representatives, as well as with the general public in matters of spatial planning to promote more sustainable and inclusive urban regions. This is not to argue that accessibility should replace other well-proven methods. However, as an additional approach, it provides views that might foster more informed debate and ultimately better decisions.

Based on the data collection and accessibility mapping exercise described here, several key strengths and limitations can be identified. Data availability and quality is the major area of concern. However, specific institutional and business issues relating to the freight transport sector are also important, as are questions regarding the context in which accessibility measures are being used. Each of these areas will be discussed below.

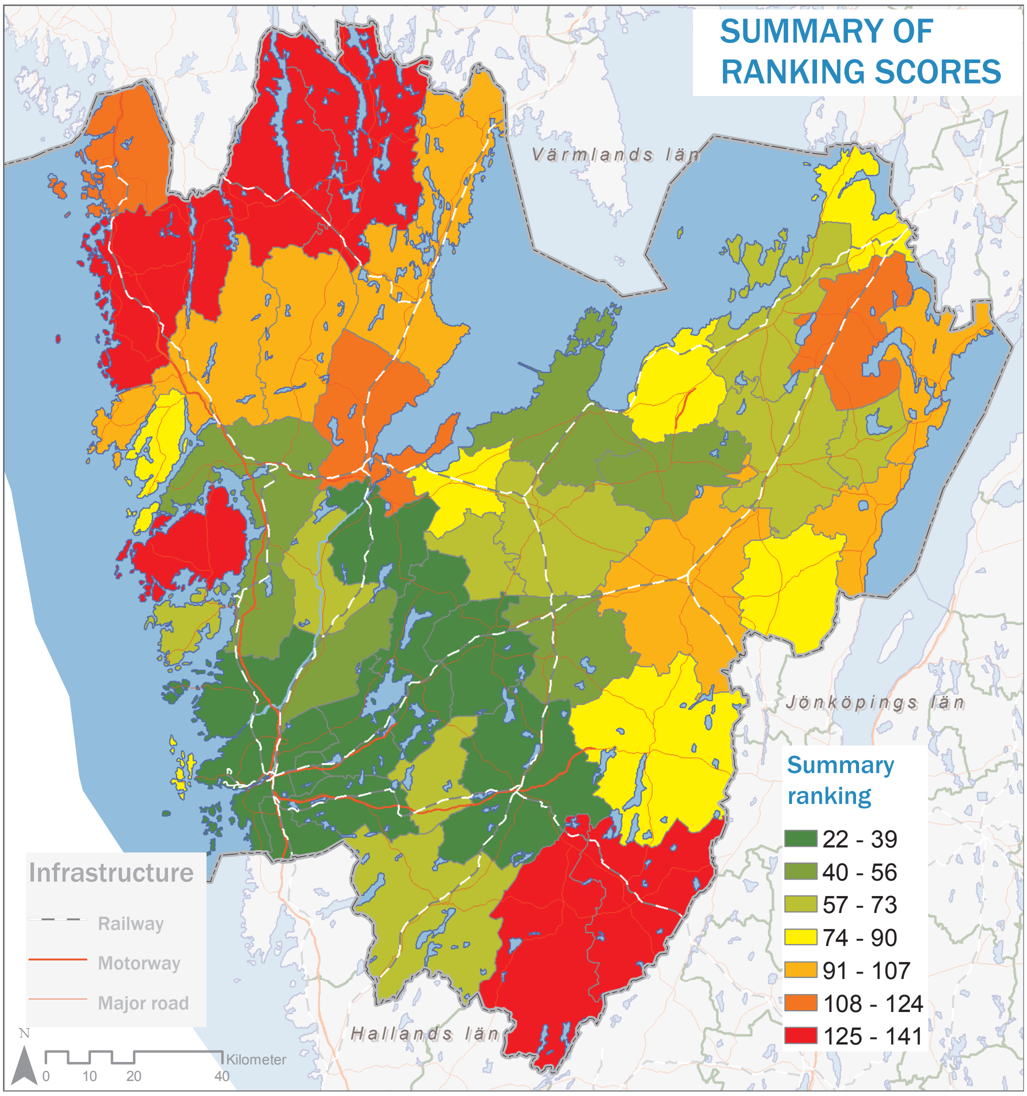


Figure 10: Illustration of the summary ranking scores per municipality

The single most pressing concern is data access and quality. Constructing useful measures of accessibility for freight transport calls for information about points of supply and demand and their locations. Since the industry encompasses a vast array of goods and destinations, it is insufficient to identify just any terminal or any customer. A basic need is business sector information in order to identify customer segments and terminal functions and to discern service levels.

The terminal data for the example maps shown here were produced in two stages. First, an automated search script scanned the official national company register (WSP 2013). Second, this was followed by a time-consuming manual check by the authors, resulting in a 25% increase in the number of terminals, many reclassifications, and the identification of multiple counting of the same terminal used by different companies. We estimate that officially available firm-level data are insufficiently reliable for accessibility analysis without considerable manual updating, possibly supported by enhanced models for the automated searching of business registers.

Specific surveys offer a potential way forward. In the case of freight transport, there is a lack of open statistical data on how goods actually flow, who is transporting goods, and what kinds of goods these are. This information is central to building more advanced models of accessibility in which the "travel behaviour" of the goods is taken into account. The assumption that the nearest terminal is the one selected is unrealistic, as indicated by current projects collecting detailed information on freight flows (see Dablanc 2013, Giuliano et al. 2015).

The question of terminal access is related partly to data availability and partly to logistics chain strategies. Many transport terminals are integrated into company-internal flows of material. For example, large retail chains or manufacturing firms employ their own global high-precision logistics systems with large warehouses. These are unavailable to local firms and should therefore not be part of an accessibility analysis of the general potentials of a specific area. From a local land-use planning perspective, however, these activities contribute to both employment and negative environmental effects.

The resolution of the maps in this study is very high. Results are presented using 500-metre cells across the region, which has proven very useful for planning problems at the local and regional scales (Elldér, Larsson 2011). However, this resolution is too high for following up policies and investments at the national level: technical limitations restrict the handling of large datasets, and complementary data are lacking at such a detailed level.

Finally, we would like to emphasize the user side of accessibility measures. So far most of the discussion has been from the producer side, focusing on measures, data, and output. However, if a non-expert user such as a planner or politician cannot fully understand the maps and measures, there is a significant communication problem that seriously limits the usefulness of these tools (te Brömmelstroet 2010). This exemplifies a wider issue regarding the sometimes contradictory relationship between scientifically sound models and the usefulness of their results in non-expert environments. The desire to model reality has often resulted in attempts to create big models that quickly become so complex that their outcomes cannot be used without expert knowledge (Lee 1973, Geertman, Stillwell 2004). This is not to say that modelling should be abandoned, but we would like to emphasize the user and the context in which results are implemented.

6 Conclusion

The use of accessibility measures to support policy and planning for freight transport and urban development offers a promising and long-needed potential for cross-sectoral integration between the fields of transport and land use. However, to fully exploit this potential several barriers must be overcome. This paper has identified a few problem areas, such as data availability and quality as well as the complexity and communicability of measures. Though these are definitely challenges in themselves, they are technical and relatively specific and can be addressed with a carefully devised research programme. Instead, we would like to highlight the wider and more complex policy and planning context for the application of accessibility analysis, especially potential tensions between actors and interests in the freight transport sector and its relations with society in general and urban development issues in particular.

While personal transport and accessibility can be related directly to individuals and their everyday activities as well as to society and planning on multiple scales, freight transport is primarily an activity organized by the private business sector. This has important implications for the usefulness of the accessibility concept and the methods of implementing it. Figure 11 illustrates this as two routes from an overall national policy objective towards more detailed planning objectives and implications in the private business and public realms. We are aware that the distinction between private and public is far from straightforward and binary; however, for reasons of clarity, we will pursue this division in the following discussion.

The point of departure is the black box in the centre of the figure containing the formulation of the overall national transport policy objective. In transport and land-use planning, this has predominantly been treated as the downward, stepwise application of national policy to regional and local levels of public planning. The more detailed the level, the further from the national goal and the more important the local context is, or should be, in terms of how to interpret and implement policies in a democratic and inclusive way. In contrast, if we follow the upward route, policies play a much more marginal role in the everyday running of businesses. Their main objective, emanating from commercial,

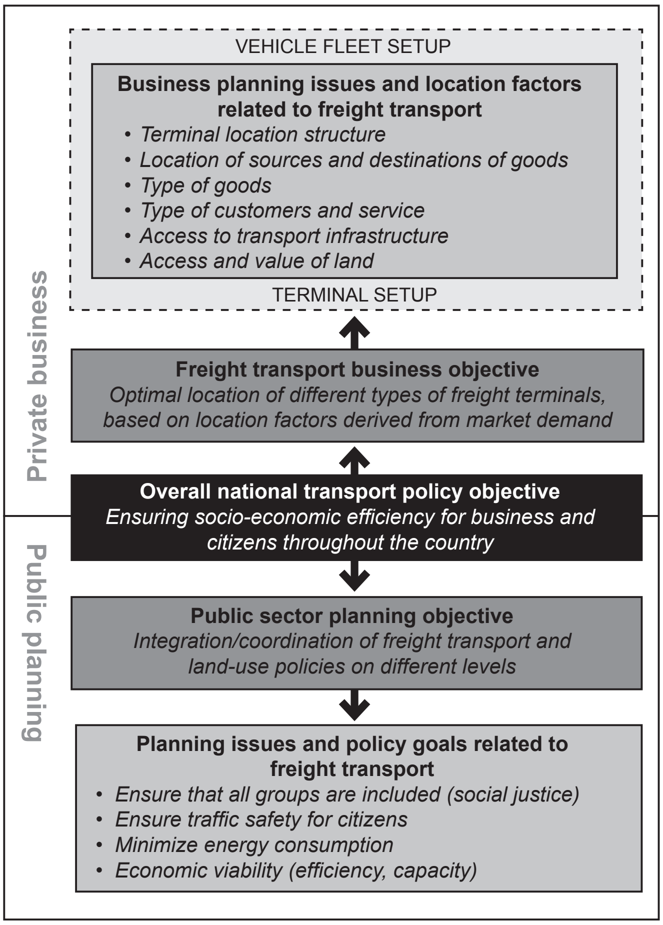


Figure 11: How the national transport policy objective can be implemented in the business and public planning spheres

market pressures, is to provide services that customers are prepared to pay for in order to earn a profit and meet the company's and shareholders' financial targets.

We might say that from the national policy objective, the scope and impact will widen downwards in the public planning sphere and narrow upwards in the business context. This is not to say that the two spheres are unconnected or exist in two parallel worlds, as urban planning is constantly interacting with freight transport and vice versa. However, accessibility analysis and accessibility questions affect the two spheres in different ways.

Imagine, for example, an expanding international port located in an expanding urban area. What would be the relevant accessibility question to ask in order to follow up the national goal? From a business competitiveness perspective, it might be interesting to know whether companies can access more and better logistics services over time in the region or whether more companies can reach the port within the same time as an effect of new infrastructure investment. For an urban planner or local policy maker in the same port city, the question of the accessibility of jobs, schools, and recreation for citizens might be at the top of the agenda. The same infrastructure that supports increased terminal and port accessibility could create physical barriers, including noise and pollution, that decrease the accessibility of other amenities for citizens.

The combination of transport and land-use components is the main strength of the accessibility concept. However, for its productive use we must be careful about its implementation context. The above example illustrates how different groups might have conflicting interests in the same space. To better measure and follow up urban transport land-use policies and potential conflicts, we must develop our knowledge, mainly about the private business sphere. The "travel behaviour" of goods and locational decision making in the freight sector are keys to implementing the accessibility concept as a long-needed integration tool in order to better understand the interplay between freight transport and its urban context.

References

Allen J, Browne M (2008) Review of survey techniques used in urban freight studies. Transport studies group, University of Westminster, London

Banister D (2005) *Unsustainable Transport: City Transport in the New Century*. Routledge, London

Banister D (2008) The sustainable mobility paradigm. *Transport Policy* 15: 73–80. CrossRef.

Barysiene J, Batarliene N, Bazaras D, Ciziuniene K, Griskeviciene D, Griskevicius AJ, Lazauskas J, Maciulis A, Palsaitis R, Vasiliauskas AV, Vasiliene-Vasiliauskiene V (2015) Analysis of the current logistics and transport challenge in the context of the changing environment. *Transport* 30: 233–241

Berglund S (2001) GIS in transport modelling. Phd dissertation. Department of Infrastructure, Royal Institute of Technology, Stockholm, Sweden

Bertolini L (2005) Cities and transport: Exploring the need for new planning approaches. In: Albrechts L, Mandelbaum S (eds), *The network society: A new context for planning*. Routledge, Oxford

Breheny M (1978) The measurement of spatial opportunity in strategic planning. *Regional Studies* 12: 463–479. CrossRef.

Dablanc L (2013) Commercial goods transport, Paris, France. Case study prepared for global report on human settlements 2013 http://www.unhabitat.org/grhs/2013

Dablanc L, Ross C (2012) Atlanta: a mega logistics center in the Piedmont Atlantic Megaregion. *Journal of Transport Geography* 24: 432–42. CrossRef.

David PD (2015) Factors affecting municipal land use cooperation. *Land Use Policy* 42: 170–182. CrossRef.

de Dios Ortúzar J, Willumsen LG (2011) *Modelling Transport*. John Wiley & Sons, Chichester. CrossRef.

Elldér E, Ernstson U, Fransson U, Larsson A (2012) Analysverktyg för tillgänglighetsberäkning med bil och kollektivtrafik i Västra Götaland. Slutrapport. occasional papers 2012:1. Department of Human and Economic Geography, University of Göteborg, Göteborg, Sweden

Elldér E, Gil Solá A, Larsson A (2012) Spatial inequality and workplace accessibility. *Environment and Planning A* 44: 2295–2297

Elldér E, Larsson A (2011) Improving regional transport accessibility planning. A GIS-based methodology based on micro-level register data. Regional Studies International Conference. 18-20 April 2011. Newcastle upon Tyne, UK

Evans I (1977) The selection of class intervals. *Transactions of the Institute of British Geographers* 2[1]: 98–124. CrossRef.

Ferreira A, Beukers E, te Brömmelstroet M (2012) Accessibility is gold, mobility is not: A proposal for the improvement of transport-related Dutch CBA. *Environment and Planning B* 39: 683–697

Geertman S, Stillwell J (2004) Planning support systems: an inventory of current practice. *Computers, Environment and Urban Systems* 28: 291–310. CrossRef.

Geurs K, Ritsema van Eck J (2001) Accessibility measures: review and applications. Rivm report 408505 006. Netherlands National Institute of Public Health and the Environment, Bilthoven, The Netherlands

Geurs K, van Wee B (2004) Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography* 12: 127–140. CrossRef.

Giuliano G, Dablanc L, Rodrigue J, Seo S (2015) Building the database. Final report project no.: 15-1A. metroFreight Center of Excellence, University of Southern California, Los Angeles, California, USA

Handy S (2002) Accessibility- vs. mobility-enhancing strategies for addressing automobile dependence in the U.S. European Conference of Ministers of Transport, Brussels

Hansen WG (1959) How accessibility shapes land use. *Journal of the American Institute of Planners* 25: 73–76. CrossRef.

Haugen K (2012) The accessibility paradox. Gerum 2012:1. Department of Geography and Economic History,University of Umeå, Umeå, Sweden

Haugen K, Holm E, Strömgren M, Vilhelmson B, Westin K (2012) Proximity, accessibility and choice: A matter of taste or condition? *Papers in Regional Science* 91: 65–84. CrossRef.

Haugen K, Vilhelmson B (2013) The divergent role of spatial access: the changing supply and location of service amenities and service travel distance in Sweden. *Transportation Research A* 49: 10–20. CrossRef.

Hesse M (2008) *The City as a Terminal. The Urban Context of Logistics and Freight Transport*. Ashgate, Aldershot

Hägerstrand T (1970) What about people in regional science? *Regional Science Association Papers* XXIV: 7–21. CrossRef.

Ingram DR (1971) The concept of accessibility: A search for an operational form. *Regional Studies* 5: 101–107. CrossRef.

Jiang B (2013) Head/tail breaks: A new classification scheme for data with heavy-tailed distribution. *The Professional Geographer* 65[3]: 482–494. CrossRef.

La Paix L, Geurs K (2016) Train station access and train use: a joint stated and revealed preference choice modelling study. In: Geurs K, Patuelli R, Ponce Dentinho T (eds), *Accessibility, equity and efficiency: Challenges for transport and public services*. Edward Elgar Publishing, Cheltenham. CrossRef.

Larsson A, Elldér E, Vilhelmson B (2014) Accessibility atlas to analyse regional accessibility to labour in the food sector. In: Brömmelstroet M, Silva C, Bertolini L (eds), *Assessing usability of accessibility instruments*. Report 2 from the COST Action TU1002, European Science Foundation and the COST Office, Brussels

Lee DBJ (1973) A requiem for large scale models. *Journal of the American Institute of Planners* 39: 163–178. CrossRef.

Lindström A (2007) Territorial governance in transition. *Regional and Federal Studies* 17: 499–508. CrossRef.

Lucas K (2012) Transport and social exclusion: Where are we now? *Transport Policy* 20: 105–113. CrossRef.

McKinnon A (2009) The present and future land requirements of logistics activities. *Land Use Policy* 26: 293–301. CrossRef.

Melo MT, Nickel S, Saldanha-da Gama F (2009) Facility location and supply chain management – A review. *European Journal of Operational Research* 196: 401–412. CrossRef.

Ministry of Enterprise and Innovation (2011) Framtidens citylogistik. Rapport från arbetsgruppen för citylogistik inom Logistikforum. Näringsdepartementet, Stockholm, Sweden

Newman P, Kenworthy J (2015) *The End of Automobile Dependence. How Cities are Moving Beyond Car-Based Planning*. Island Press, Washington DC

Olsson J, Larsson A (2016) Accessibility to terminals in Västra Götland – A pilot study. PM 2016:9, Transport Analysis, Stockholm (in Swedish) http://www.trafa.se/-globalassets/pm/pm-2016_9-tillganglighet-till-terminaler-i-vastra-gotaland--en-pilot-studie.pdf

Olsson J, Woxenius J (2012) Location of freight consolidation centres serving the city and its surroundings. *Procedia – Social and Behavioral Sciences* 39: 293–306. CrossRef.

Olsson J, Woxenius J (2014) Localisation of freight consolidation centres serving small road hauliers in a wider urban area: barriers for more efficient freight deliveries in Gothenburg. *Journal of Transport Geography* 34: 25–33. CrossRef.

Páez A, Scott DM, Morency C (2012) Measuring accessibility: positive and normative implementations of various accessibility indicators. *Journal of Transport Geography* 25: 141–153. CrossRef.

Rodrigue J, Comtois C, Slack B (2013) *The Geography of Transport Systems*. Routledge, Abingdon

Sakai T, Kawamura K, Hyodo T (2015) Locational dynamics of logistics facilities: Evidence from Tokyo. *Journal of Transport Geography* 46: 10–19. CrossRef.

Sheffi Y (2012) *Logistics Clusters. Delivering Value and Delivering Growth*. MIT Press, Cambridge

SOU – Statens Offentliga Utredningar (2008) Vägen till ett energieffektivare Sverige. Slutbetänkande av Energieffektiviseringsutredningen, SPU 200:110

Straatemeier T (2008) How to plan for regional accessibility? *Transport Policy* 15: 127–137. CrossRef.

Swedish Transport Administration (2013) Godskartläggning i VG. Rapport 2013:151. Trafikverket, Borlänge

te Brömmelstroet M (2010) Equip the warrior instead of manning the equipment: Land use and transport planning support in the Netherlands. *Journal of Transport and Land Use* 3: 25–41

Transport Analysis (2013a) Indikatorer för en transportpolitisk måluppföljning – hur tillgänglighet påverkar konkurrens- och utvecklingskraft. Rapport 2013:2. Trafikanalys, Stockholm, Sweden

Transport Analysis (2013b) Indikatorer för en transportpolitisk måluppföljning – hur tillgänglighet påverkar konkurrens- och utvecklingskraft. Trafikanalys: Stockholm, Sweden

Transport Analysis (2013c) Transportarbete 1950-2012. *Trafikanalys: Stockholm, Sweden. http://www.trafa.se/sv/Statistik/Transportarbete/*

WSP (2013) Methodology for the identification of freight terminals in Sweden. Internal report (in Swedish). WSP, Stockholm, Sweden

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 1, 2017, 93–107
DOI: 10.18335/region.v4i1.158

journal homepage: region.ersa.org

Logistics sprawl in monocentric and polycentric metropolitan areas: the cases of Paris, France, and the Randstad, the Netherlands

Adeline Heitz[1], Laetitia Dablanc[2], Lorant A. Tavasszy[3]

[1] University of Paris East, Paris, France (email: adeline.heitz@ifsttar.fr)
[2] IFSTTAR, Paris, France (email: laetitia.dablanc@ifsttar.fr)
[3] Delft University of Technology, Delft, The Netherlands (email: l.a.tavasszy@tudelft.nl)

Received: 5 September 2016/Accepted: 13 April 2017

Abstract. The phenomenon of urban sprawl has been studied extensively. Most research so far has focused on residential settlements. A growing number of studies have addressed industrial deconcentration. Our focus in this paper is on logistics sprawl, i.e. the growth and suburban relocation of warehousing activities. Specifically, we investigate the difference in logistics sprawl between monocentric and polycentric systems of cities. The literature suggests that logistics activities, like residential settlements, will gradually move to suburbs as land prices increase in central areas. As research on logistics geography has mostly focused on monocentric systems, the question is whether this also applies to polycentric systems. We compare two cases, the Paris region in France, representative of a monocentric urban development, and the Dutch Randstad area as a polycentric case. We use regional statistics on warehouse settlements in both regions for a descriptive analysis of changes since the mid-2000s to derive metrics for concentration. The cases show different patterns of change. In contrast to Paris, logistics activities within the Randstad have intensified in dense areas. We explore the reasons that may explain this difference and conclude that urban structure, spatial planning policies, and the freight hub quality of a region are factors of influence.

1 Introduction

The recent growth of logistics activities including warehousing has increased their importance in metropolitan areas (Bowen 2008). Logistics activities are sprawling outwards and warehouses are leaving dense areas and relocating in peripheral ones. Logistics sprawl is the subject of a growing volume of research. North American metropolitan areas such as Chicago (Cidell 2010), Atlanta (Dablanc, Ross 2012) or Toronto (Woudsma et al. 2015) have been covered by pioneering work on logistics sprawl. Several studies have examined the case of Paris (Dablanc, Rakotonarivo 2010, Raimbault, Bahoken 2014). With the exception of Paris there have been few case studies in Europe.

Yet European metropolitan areas, far from forming a homogeneous whole, are characterized by long-established and variable urban forms. The literature (see below) distinguishes between two main types of urban region structure: monocentric and polycentric, which are actually the two extremes of a continuum of urban regions. Europe has a large number of monocentric urban regions such as Paris or London, but also polycentric urban regions like the Randstad region (Davoudi 2003). Polycentric and monocentric urban regions each provide a specific urban context for the development of economic activities. Monocentric urban regions are characterized by their large and growing population and the fact that economic and leadership activities are concentrated in one metropolitan area. These factors mean that they also have a considerable capacity for growth, so, in spatial terms, they expand physically. The dense, uninterrupted, parts of a large city are easily identifiable. However, urbanized areas of its outer suburbs, which are scattered over distances of several tens of kilometers from the center, are difficult to distinguish from rural areas. The urban area includes the central conurbation and its suburban rings, which itself contain secondary centers of various sizes. This urban region has the appearance of a "macroform" (Allain 2004). Polycentric urban regions are, on the opposite, made of several urban centers, which polarize population and activities, and which are linked to each other by flows (freight, people, information) and cooperation or interdependent bounds of various types. Cities within a polycentric urban region seem to have merged in a functional and morphological entity: a large and dispersed regional urban system. Polycentric regions are characterized by the fact that residential and economic functions are distributed within the metropolis between clearly separated specialized centers of employment (Berroir et al. 2008). The residential function is organized around these centers, forming high-density clusters.

In this paper, we examine the impact of the metropolitan structure on the development of logistics spatial dynamics. We have chosen two archetypal urban regions, Paris and the Randstad, in order to take a different view on the location's factors traditionally observed in the literature. By using the same method and the same kind of data for both areas, we want to contribute to the general debate on the location's factors for logistics facilities. By observing the evolution of the location of logistics facilities in metropolitan areas which are located either in a monocentric urban region or in a polycentric urban region, we want to test the recurrence of the logistics deconcentration and learn from potential differences. This paper also contributes to establishing the distinction between "logistics sprawl" and "logistics suburbanization".

The remainder of the paper is built up as follows. Section 2 provides a literature review on logistics deconcentration, logistics sprawl and the main drivers that usually explain this dynamic. Section 3 explains the data used for the analysis and the results of the comparison of the two regions and provides an interpretation of the results regarding contrasting forms of logistics urbanization in the two regions. Section 4 provides an analysis of new drivers explaining the deconcentration process in those metropolitan areas, including public policies and their potential impact on logistics spatial patterns. We conclude the paper in Section 5.

2 Logistics sprawl and logistics suburbanization

2.1 Logistics sprawl in metropolitan areas: a state of the art

Over the last thirty years, a process of metropolization has further increased urban growth and urban sprawl. Metropolitan areas have spread out because of a lack of available space in the dense parts of the conurbation (Nicot 1996) to accommodate this growth. Urban deconcentration is the outcome of the gradual saturation of the most central areas and increases in the price of land because of its relative scarcity. Logistics deconcentration, or sprawl, means an increasing number of logistics facilities in peripheral areas and their dispersion. It describes the decreasing number of logistics activities, such as warehouses, in urban centers, and their redistribution in the periphery. This is what is known as "centrifugal growth" in the typology developed by Champion (2001). What occurs is that secondary centers emerge in peripheral areas where property prices and traffic congestion are lower. These secondary centers are easily accessible thanks to high speed transport infrastructure. According to the typology developed by Le Nechet, Aguiléra (2012), Paris is a type of "pyramid" metropolis, that is to say an isotropic space in which land market forces theoretically lead to an exponential decrease in population density with distance from the employment center.

Several factors explain logistics sprawl. First, the nature of supply chain management has shifted concomitantly from "supply-push" to "demand-pull" systems (Lasserre 2004).

In the former, firms "pushed" their output into distribution channels based on demand forecasts, relying on storage places close to production sites. In the latter, firms gear their production in response to real-time information about what consumers are buying, and use distribution centers closer to consumer markets (Bowen 2008). Changes in supply chain management impacted the location of logistics facilities. Moreover, historically, the majority of these storage buildings were smaller and located in inner urban areas in the proximity of industrial areas, rail terminals, and docklands (Dablanc, Rakotonarivo 2010, Cidell 2010). Changes in the location of transportation terminals also influenced the location. Because of an increasing size of facilities, large and cheap available land parcels were more likely found in suburbs and exurbs.

In a recent literature review paper, Aljohania, Thompson (2016) identifies studies showing that "land use control and the exclusion of freight in urban planning have influenced a relocation of logistics facilities from inner urban areas to suburban areas, as affordable industrial land was no longer available for logistics companies". Nicot (1996) concludes that the lack of industrial land in inner urban areas is a primary driver in the relocation of logistics facilities away from urban cores. Merenne-Schoumaker (2008) identifies several indicators that guide the choice of location of logistics activities and warehousing establishments. First of all, land availability appears to be decisive. Distribution centers tend to require fewer and larger single-story facilities from 10,000 to 100,000 m2 ("XXL warehouses") (Hesse 2004, Cidell 2010). Consequently, suburban and exurban areas offer more affordable locations for logistics facilities. In a monocentric urban area, the price of land decreases from the center of the metropolitan area to the periphery. In polycentric urban areas, the price of land decreases for each center to each periphery, but the decrease is limited by the proximity to each metropolitan area. Suburban areas appear to be favored land for logistics facilities' development. In order to compensate for the remoteness, logistics facilities rely on accessibility and choose to locate close to the main roads and highways, which ensure access to labor markets and consumption areas. The last factor of location, identified by Merenne-Schoumaker and also underlined by Aljohania, Thompson (2016), is the role of public policies in land use. By regulating, funding, or planning dedicated zones for logistics, public policies encourage the location of logistics facilities in determined areas. The location of logistics facilities is not only driven by market forces and land opportunities.

2.2 From logistics sprawl to logistics suburbanization

Land availability appears to be decisive. Consequently, suburban and exurban areas offer more affordable locations for logistics facilities. Suburban areas appear to be favored land for logistics facilities' development. In the case of the Paris region, Dablanc, Rakotonarivo (2010) showed that the parcel industry has moved to the periphery as a result of centrifugal forces. Paris infrastructure heritage initially favored the concentration of logistics and freight transport in dense central areas. Then some of these logistics activities, including distribution centers (Raimbault 2015), were transferred to peripheral areas.

In this paper, we make a distinction between 'logistics sprawl' and 'logistics suburbanization.' Logistics suburbanization has an additional qualitative meaning to the mere relative increase in the number of warehouses in areas further away from the city center. Logistics suburbanization describes the transformation of low density suburbs and the irruption/integration of logistics in the outskirts of the metropolitan area. Suburbanization can be defined as the urbanization of peripheral areas (generally with low density settlements). Spaces with lower densities provide the most attractive location for these warehouses, especially as modern warehouses are larger and in requirement of large land parcels. This development of logistics in the outskirts not only contributes to the expansion of metropolitan areas, but also reshapes the suburban areas that accommodate such activities. Logistics suburbanization corresponds to a "third phase" of logistics development (Frémont 2015). Logistics functions are discharged to external centers in the urban periphery where available space for large warehouses can meet current demand. This may present a problem not only for their spatial optimization but also for land use efficiency. Indeed, the development of logistics areas in the metropolitan fringes can promote a process of urban disintegration (Hall, Hesse 2013).

Logistics suburbanization depends on/results in a high level of goods transport. The concentration of logistics activities in the suburbs redefines the functions of both dense and suburban areas. Suburban areas become supporting territories, bases or intermediate points for the flow of goods, and often the last place before the final distribution of goods. The relationship between central and peripheral areas is linked to the flow of goods from a suburban warehouse into dense city areas.

Hall, Hesse (2013) pointed out in the conclusion of their study that the deconcentration of logistics activities as a suburban development is a "surprisingly common" phenomenon in many cities, presenting a challenge for public authorities. It seems appropriate to put this conclusion into perspective by comparing the urban structures of two very different cities. Most of the factors of location, which have been described previously, have been analyzed in the case of a single metropolitan area in a monocentric urban region. By measuring logistics sprawl in different metropolitan areas located in two different urban regions, one monocentric and one polycentric, we want to test the robustness of those factors and emphasize the importance of the structure of the urban region.

3 Logistics sprawl in the Paris and Randstad metropolitan areas

3.1 Method and data

The urban region of Paris (the Ile-de-France region), with 12 million inhabitants on 12,000 km^2, consists of nearly 1,300 municipalities. Ile-de-France is the largest consumer market in France, which demands an efficient logistics organization. The region, which represents 25% of the country's population, and 30% of its GDP, contains approximately 20% of France's total warehousing space. The transportation and logistics sectors account for almost 10% of employment in the region (about 400,000 jobs). Between 2000 and 2012, the Paris region experienced a 33% increase in the number of its warehousing facilities.

The Randstad region in the Netherlands, with 7.1 million inhabitants, includes 183 municipalities over 8,300 km^2. It is not an administrative entity, but the product of collective construction, gradually becoming included in the Dutch planning process. The term, literally meaning "border city", was coined in 1930 by a pilot who had noticed the peculiar urban form that extended from Rotterdam to Utrecht, surrounding a vast natural area ("Groene Hart" or Green Heart) (Kühn 2003). This term was adopted by public authorities in planning documents in the 1950s and has been steadily gaining currency since (Burke 1966). In 1966, Peter Hall (2008) described the Randstad. His notion of "metropolis" is still being debated today, for example in the most recent national planning document "Randstad Strategic Agenda 2040" (van der Burg, Vink 2008). The region is made up of four provinces (Noord Holland, Zuid Holland, Flevoland and Utrecht), which contain some of the largest cities in the Netherlands (Amsterdam, Rotterdam, The Hague and Utrecht). The Randstad represents 45% of the population of the Netherlands on 25% of its land and is the main gateway as well as goods consumption area in the Netherlands, in which respect it is comparable with metropolitan Paris. The Randstad region is a polycentric urban region clearly identified as a cultural, social, or political entity, which makes it an interesting case study.

This study aims to present comparable data for the two metropolitan areas. We used two databases: 'Local Knowledge of the Productive System' (CLAP) provided by INSEE (National Statistics Institute) for the French data and the Lisa database for the Dutch data. These two databases have the advantage that both use the NACE (Statistical Classification of Economic Activities in the European Community), so we can use the same categories for our comparison. We chose to use Category 52.1 "Storage". This category applies to activities that require a logistics building, namely a warehouse. In this category, the floor size of warehouses is not specified. According to DRIEA (2009), about 20% of the warehouses in the Paris region are less than 5,000m^2. Our databases also lack precision because they do not include all logistics buildings of the "warehouse" type. This is because the NACE classification only considers the principal purpose of a building, so some warehouses that serve the parcel industry or the distribution sector may not be recognized. The classification used by NACE and by extension the CLAP and Lisa databases have many shortcomings, which make it difficult to obtain an accurate estimate of the number of establishments and their location. Because of these limitations, the final number of logistics facilities presented in this study is only an estimate of the total.

We used the aggregated number of warehouses in municipalities for the years 2004 and 2012 for Ile-de-France and for the years 2007 and 2013 for the Randstad. Although the dates are not the same, as we did not have access to the exact same year, they are quite close from each other. Both allow us to understand how the location of these facilities has changed since the mid-2000s. Another comparability problem for these data is the great difference in the size of the municipalities in the two countries. In Ile-de-France the average size of municipalities is 10 km^2 with a population density in each municipality of 9,000 inhabitants per km^2 while in the Netherlands it is 86 km^2 with a density of about 1,000 inhabitants per km^2. To overcome this problem, we chose to use statistics for 'cantons' (an electoral jurisdiction) in Ile-de-France. The most rural municipalities have thus been grouped together, while those in the densest cities have remained separate. This particular division of the territory allows us to have a similar population in each division and just 286 statistical entities in Ile-de-France, which makes it possible to make a comparison between the two areas. This limitation does not fully allow us to compare the population densities and integration of warehouses in the metropolitan structure in absolute terms. We will make relative comparisons in view of the different situation in the two countries.

We have used a basic but robust spatial analysis indicator known as the centrographic method. This allows us to measure the changes in the distribution of a statistical population (in our case, warehouses) in space and over time, through the use of several indicators. We have calculated the mean distance of the warehouses in the two regions and in each province of the Randstad (we used the provincial unit for ease of data access and processing). Each Randstad province contains a rather well defined monocentric urban structure (with the exception of Flevoland, with no major city). To measure sprawl, we used the change in the average distance of terminals to their center of gravity. We used the calculation method that is included in the ArcGIS software (Mitchell 2005). The data on warehouses is aggregated for each province of the Randstad region and for each canton for the metro area of Paris. We weighted each province and canton by the number of warehouses.

3.2 Measuring logistics sprawl in the Paris and Randstad regions

The first observation of interest is that at macro-level scale, the two regions have experienced a different situation regarding logistics sprawl, with a deconcentration of logistics facilities in the Paris region (+5 km) and a contraction in the Randstad region. In the same time we observed an increasing number of warehouses in the Paris region (+33%) and stagnation in the Randstad (-1%). Not only are spatial dynamics different in the two but also the state of the logistics real estate market.

The Paris metropolitan area experienced a considerable degree of warehouse deconcentration between 2004 and 2012, since the average distance of the warehouses from their center of gravity increased by 4.1 km between these two years. This finding backs up the conclusions reached by an analysis of the parcel industry from the 1970s to 2010 (Dablanc, Rakotonarivo 2010) and of all warehouses during the last fifteen years (Heitz, Dablanc 2015). Our new analysis shows that logistics sprawl has continued in the Paris region (Ile-de-France) and the Paris metropolitan area ('Grand Paris') between 2010 and 2012.

The second observation is the heterogeneous nature of logistics spatial dynamics in the different provinces of the Randstad. The provinces of Noord Holland and Zuid Holland underwent contraction around the center of gravity. Indeed, the average distance from the center of gravity has fallen by -2 km in the province of Noord Holland and by -1 km in that of Zuid Holland. Although this is a relatively small decrease, it shows the existence of a dynamic in both provinces during the last years. Noord Holland is the province of the City of Amsterdam, which has attracted many service activities. The province of Zuid Holland is the largest of the Randstad and contains the conurbation between Rotterdam and The Hague, with several centers.

The observed geographic patterns of change may be a result of both centripetal forces

Table 1: Spatial indicators

	Regions		Metro areas				
	Paris Region (Ile-de-France)	Randstad	Metro area of Paris ('Grand Paris')	Noord Holland (Amsterdam)	Zuid Holland (Rotterdam)	Flevoland	Utrecht
Time period	2004-2012	2007-2013	2004-2012	2007-2013	2007-2013	2007-2013	2007-2013
Area (km^2)	12,012	8,357	657	2670	2818	1419	1450
Population (million)	11.9	8.5	4.5	2.7	3.6	0.4	1.2
Number of warehouses (most recent year)	955	583	441	278	185	59	61
Number of warehouses (most ancient year)	713	589	388	318	168	60	43
Change in number of warehouses (%)	+34	-1	+14	-12	+10	-2	+42
Logistics sprawl indicator (km)	5	–	4.1	-2	-1	3.3	0.5

and centrifugal ones at the Randstad level. Logistics sprawl at the Randstad regional level may translate into concentration or deconcentration at the local scale (for each metropolitan area). The Randstad region experienced a dispersion of logistics activities in the 1970s when they moved from clusters to peripheral regions within the Netherlands (Davydenko et al. 2013). Nevertheless, our analysis measures these patterns in more detail and, also seems to suggest a coordinated move of centers of gravity towards the "Groene Hart", i.e. sprawl into the heart of the ring-shaped Randstad area, along the direction of the main highways connecting the cities. The provinces of Zuid Holland and Noord Holland have a different situation regarding the evolution of the number of warehouses. In the province of Zuid Holland the number of warehouses has increased by 10% while the number of warehouses has fallen in the province of Noord Holland by 12%. The absence of logistics growth and of logistics sprawl in Noord Holland means that Amsterdam metropolitan area as a logistics cluster is shrinking. Meanwhile, logistics facilities are growing inside the dense part of the Rotterdam metro area.

The deconcentration of warehouses is a confirmed pattern in the provinces of Utrecht and Flevoland. The province of Utrecht is small and polarized by the city of Utrecht. Flevoland province is slightly different from the others: it does not include major cities such as Amsterdam and Rotterdam nor major transport infrastructures, and in historical terms it is a recent creation. It cannot therefore be considered as an urban center in the same way as the others, but it is the subject of many proactive development policies, particularly for its new town of Almere. The deconcentration of warehouses is taking place in the provinces of Flevoland (+3.3 km) and Utrecht (+0.5 km), although the intensity of the process is not the same, being relatively low in the province of Utrecht compared with Flevoland. We can hypothesize that the significant deconcentration taking place in the Flevoland province is linked to the recent nature of its development.

By using the centrographic method, we have been able to determine whether warehouses have expanded or contracted around the urban centers. While it allowed us to observe the expansion of logistics activities in metropolitan areas, it did not allow us to say whether there was a pattern of logistics suburbanization, as we defined it above. We must supplement this discussion with an analysis of the integration of warehouses in relation to the density of metropolitan areas.

3.3 Measuring logistics suburbanization

To measure logistics suburbanization, we have established a profile for each municipality for the Randstad and each statistical canton for Ile-de-France, according to two criteria: firstly population density, and secondly the number of warehouses, which to some extent reflects logistics intensity. We have used the same data as before: the number of establishments classified as "warehouses" in the CLAP data for Paris and the Lisa data for the Randstad. These municipal (or cantonal) profiles were prepared using quartile-based discretization for each of the studied variables (population density and number of logistics facilities). This gave sixteen profiles. For ease of understanding we have grouped these profiles together to form four major types (see Table 2 and Figure 1).

Once the profile of each municipality has been assigned to one of the four possible types, we have calculated which profiles are overrepresented in order to identify the dominant patterns in each metropolitan area and obtain a specific picture of logistics sprawl (Figure 1). The cantons corresponding to the profile "A" have a population density of 11,687.5 inhabitants/km^2 in the Paris region compared with 594 inhabitants/km^2 in the Randstad. The profiles at each date and the way they have changed show the pattern of logistics sprawl that has taken place in the metropolitan areas.

Figure 1 shows how logistics activities are integrated within the metropolitan areas of each region and how they have changed over time.

In the case of Paris, the overrepresented profiles were types A and D in 2004, while in 2012 it was types B and C. In recent years, more warehouses have been set up in suburban towns than in municipalities with a higher population density. We do observe logistics suburbanization in the Paris region. This observation concurs with the work of Raimbault, Bahoken (2014) on logistics suburbanization. Logistics activities are declining in the densest areas and tending to favor less dense areas in the outer suburbs or the first

Table 2: Typology of municipality/canton profiles

Profile	Type	Description
A	Logistics in dense areas	High population density High number of logistics establishments
B	Logistics in suburbs	Low population density High number of logistics establishments
C	Residential zones	High population density Low number of logistics establishments
D	Rural zones	Low population density Low number of logistics establishments

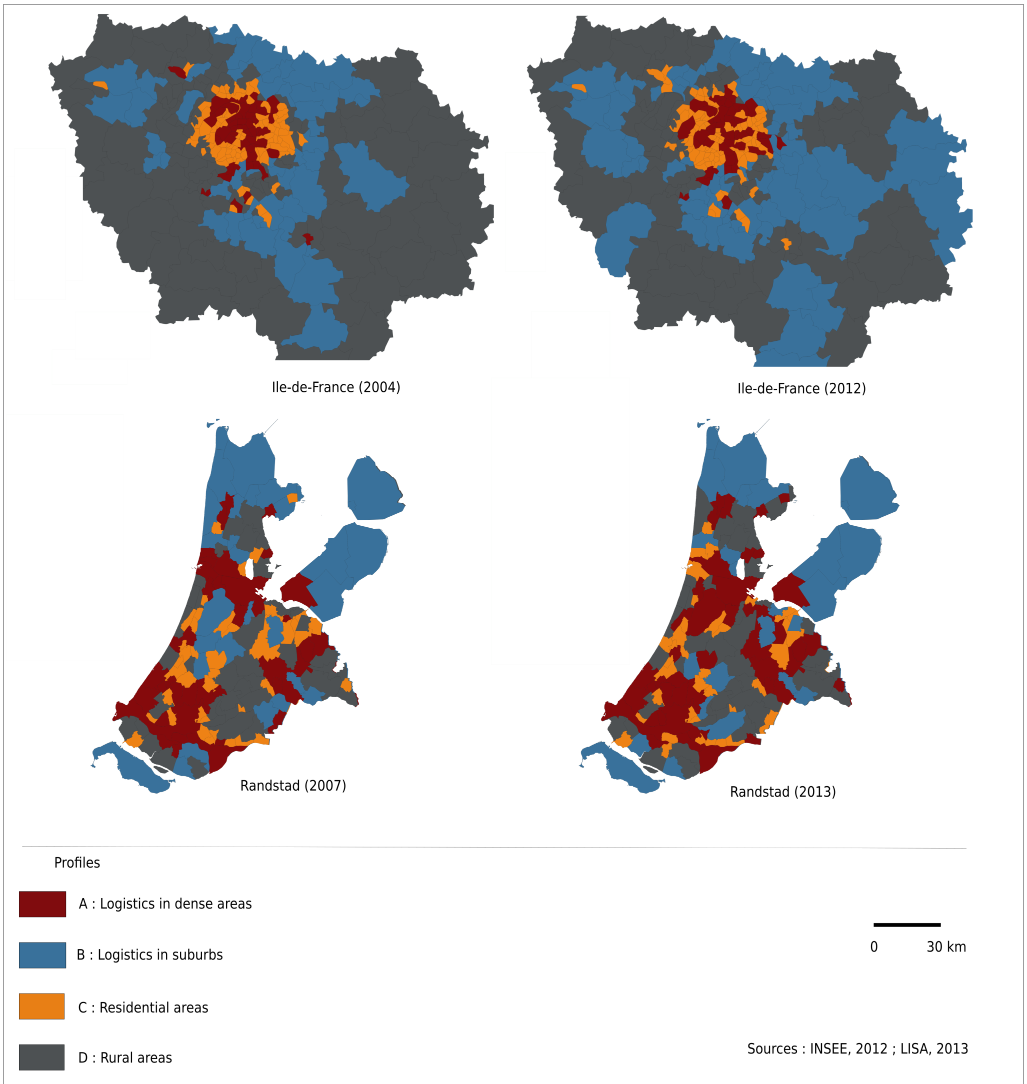


Figure 1: Comparison between local profiles in the Paris and Randstad regions

suburban ring. Logistics sprawl in metropolitan Paris has the effect of strengthening and densifying peripheral areas. The change in the location of the warehouses results not only in the expansion of the metropolitan area but also the increasing presence of logistics in the outskirts, in less densely populated areas.

In the Randstad region, municipalities with types A and D profiles dominate, both in 2007 and 2013. This means that logistics activities continue to locate in the dense areas of the region. Logistics suburbanization is much weaker than in Paris.

What is particularly striking in the analysis of the integration of warehouses in the metropolitan structure of Paris is the change in the location pattern between 2004 and 2012. Over this short period, there was a profound change. Logistics activities today seem to locate mainly in the suburbs where their numbers are growing. This analysis reveals that the warehouse location pattern goes beyond the deconcentration of logistics activities: logistics suburbanization is the dominant model. In contrast, in spite of the deconcentration of warehouses in certain urban conurbations, the Randstad exhibits weak logistics sprawl as well as logistics suburbanization, focused towards the inside of the ring-shaped region. Its dense areas have remained the prime location for logistics activities.

In recent years, modifications in the location of warehouses have generated spatial and morphological changes in cities (Frémont 2015). Overall, logistics facilities in the Paris and Randstad regions have become more dispersed. At the metropolitan scale, the Rotterdam and Amsterdam metro areas have an atypical profile and their warehouses have concentrated. They have not experienced logistics decentralization. In metropolitan Paris, the number of logistics sites has increased in peripheral areas to the detriment of the central areas, which means that logistics urbanization takes the form of logistics suburbanization. In summary, the case of the Randstad reveals the existence of a different form of logistics urbanization, logistics intensification, which we will discuss in the next section.

In section 2 we have identified different factors of relocation of logistics facilities mentioned in the literature. To explain the different forms of logistics locational patterns between the Paris and Randstad regions, we look at some of these factors in more detail.

4 Analysis of drivers explaining locational patterns: logistics suburbanization vs. logistics intensification

4.1 The provision of airports and large logistics clusters

Logistics facilities often tend to locate within logistics clusters in the fringes of metropolitan areas close to major airports and seaports or even highway networks (Hesse 2002, Woudsma et al. 2008, Cidell 2011, Allen et al. 2012). A freight transport system mainly performed by trucks contributes to logistics sprawl. In the Paris region, 90% of goods (in tonnes) are carried by road. Road accessibility is one of the most important factors of relocation of logistics facilities in the suburbs. Distance to the center is compensated by accessibility. A first movement of decentralization of transport infrastructure such as freight rail facilities in the 1960s and 1970s took away some logistics facilities from city centers to peripheral areas (Beyer 1999). In Paris, the development of logistics activities in peripheral areas is in part explained by the location of major transport nodes in these areas: Roissy and Orly airports, the river ports of Gennevilliers and Bonneuil, for example. These nodes still retain large clusters of warehouses (Heitz, Dablanc 2015). It so happens that the three major clusters in the innermost suburbs to the north (Gennevilliers), northeast (Roissy, Mitry-Mory) and south (Orly and Bonneuil) are also the main gateways for national and regional freight. These clusters are located in the peripheral areas. In the case of Paris, the location of nodal transportation infrastructure is an important factor of relocation of logistics facilities.

The Randstad region contains the port of Rotterdam, which is the largest (in terms of freight traffic) maritime port in Europe, and the Amsterdam Schiphol airport which is the fourth largest in Europe. The port of Rotterdam and its entire region are integrated in port regionalization dynamics, which means that logistics activities spread out into the region. Many activities are functionally linked to the port, and spread over a region that stretches from Dordrecht to Venlo through Tilburg (Priemus, Visser 1995) crossing the boundaries of municipalities and provinces (van der Burg, Vink 2008). However, it seems that this dynamic has been reversed in the provinces of Noord Holland and Zuid Holland in the last fifteen years. This observation allows us to appreciate how quickly the location of warehouses can change. The port of Rotterdam and Schiphol airport are the two most important logistics clusters in the Randstad (OECD 2007) and therefore naturally have a high concentration of warehouses and other logistics activities. In Zuid Holland in recent decades the port of Rotterdam has sought to limit its spatial expansion. The port of Rotterdam has decided to promote logistics development in the vicinity of port infrastructure. The Havenplan 2010 Port Development Plan emphasized the importance of limiting the space taken up by port development with the construction of the Second Maasvlakte, a polder in the North Sea near Rotterdam, and the relocation of activities (Priemus, Visser 1995). The decision to build this landfill was taken in 2004 and the facility was opened in 2013, thereby halting the expansion of the port towards the city and allowing re-urbanization of industrial areas. In anticipation of increased activity in the port, logistics activities have been redistributed around the port in recent years. The intensification of logistics reflects the existence of a new clustering of logistics activities highlighted by van den Heuvel et al. (2013).

Among the main factors for the development of logistics activities in the suburbs is a need for accessibility. Therefore, a possible explanation of logistics intensification is the severe congestion facing the area on the roads that connect the major urban centers. "Road congestion in the Netherlands is one of the most urgent problems facing the Port of Rotterdam, as it increases and decreases delivery time reliability for logistics companies that have chosen to use Rotterdam" (OECD 2007). Congestion can be a factor that limits logistics sprawl if its cost becomes excessive.

4.2 The role of public policies on logistics locational patterns

Dablanc, Raimbault (2015) argue that, despite attempts to achieve regional organization, the institutional fragmentation of metropolitan areas in France and the US lets "the laws of a dynamic land market and the most local level of government decide the location of logistics facilities" (p.301). A new logistics real estate in the outskirts of metropolitan areas reinforces the suburbanization of logistics, closely related to a lack of planning at the metropolitan level. With a long tradition of planning to address urban sprawl (Bogaerts et al. 2007), Dutch planners have de facto limited the sprawl of economic functions at a metropolitan scale. Suburban development took place in "quite a deconcentrated manner along the highways, at infrastructure nodes and on designated sites" (Bogaerts et al. 2007), but always limited by the "Groene Hart". Logistics deconcentration, as a result, appears to be quite limited in comparison with Paris. At the metropolitan scale, centripetal forces remain important, encouraged by policies focused on transport infrastructures, and logistics activities decentralize mostly along the highways. The polycentric region of the Randstad, overall, favors logistics concentration.

It is our hypothesis that this concentration is reinforced by public policies at the local and metropolitan levels. Public policies in the Netherlands have traditionally mainly focused on (1) developing the position of the Randstad as a hub for distributing goods across Europe; and (2) protecting the region's environment and contributing to the efforts against climate change. In the vision entitled "Randstad Holland in 2040" (Ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieu 2008), the government set out its plan to make the Randstad into a "sustainable, competitive top European region", considering the metropolitan region as a whole, which must become more homogeneous. With the plan "Randstad Holland in 2040", public authorities want to group together the "city regions" which are unequally developed, to form a coherent morphological and functional entity. This may help promote logistics intensification in the region. In Paris likewise, public policies have recently considered logistics sprawl. The recent regional master plan (SDRIF 2013) recommends the preservation of logistics activities in dense areas of the region and the promotion of tools to accommodate new logistics activities in the central areas. The new Paris municipal zoning code (Ville de Paris 2016) identifies more than 80 central city locations reserved for future logistics activities.

The Netherlands has a long tradition of urban planning. From the 1950s to the 2000s, a specific idea of the Randstad has predominated in the planning of the region influencing the location of economic activities. In 1958, the Randstad was defined as a group of “city regions” arranged in a horseshoe shape and separated by natural buffers including a large central green area. Urbanization is concentrated in the main urban centers and along the main roads connecting them (Rijksdienst voor het Nationale Plan 1958). The history of planning in the Randstad consists of successive policies promoting the development of compact cities and limiting urban sprawl by controlling the development of activities and the population on the region. In 1993, the Dutch Parliament adopted the “Supplement to the Fourth Policy on Territorial Planning”, or VINEX. By favoring a compact city model, the Parliament tried to regulate mobility, increase the capacity of urban infrastructure, and protect natural spaces, for example with the ABC location policy (“ABC-locatiebeleid”) (Bogaerts et al. 2007), which intended to plan the location of trip generators such as businesses, services, and industry. The location of economic activities was reorganized on the basis of the number of jobs in order to maximize the use of public transport. Activities involving fewer jobs and greater dependency on motorway access, such as warehouses, were encouraged to move to the outskirts, reserving the most densely populated areas for office type activities. The ABC location policy was then updated by policies that encourage the concentration of activities, including activities with lower density of jobs such as logistics. If the efficiency of these policies has been discussed, the Dutch planning system leaves considerable room for local spatial policy initiative and power.

Therefore, we could make the hypothesis that the concentration of warehouses around the main urban centers is a direct consequence of the desire of the Dutch government to encourage the concentration of urbanization. At the local level, the development of each urban center is a way of achieving a balance in the region. The project to develop the Randstad as an institution also aims to make a coherent area, based on a functionalist approach to space and on a division of functions between the economic centers in order to make them more specialized, intensify their activities, and improve their connectedness (van der Burg, Vink 2008). Rotterdam was identified as the main hub for the development of logistics and freight transport, Utrecht as a university center, The Hague as a political cluster that focuses on international law, and Amsterdam as the capital city, retaining the role of a multifunctional center. This distribution of functions is based on a particular concentration of economic activity around the various centers, especially around each of the major transport hubs, for example the ports of Rotterdam and Amsterdam or Schiphol airport.

We argue that logistics intensification may be partly due to the Randstad urban region’s structure (polycentrism) and partly due to the public policies of concentration and preservation of green spaces. Parallel to these, ’relative densities’ between urban centers and suburban areas may also be identified as an explanatory variable: a hypothesis can be that when the difference in density levels within a metro area is relatively minor, the availability of land or the differences in rental prices are not strong incentives to site a warehouse out of an urban center. A comparative analysis of densities and warehouses’ rental prices in our two urban regions is still required to conclude on that matter.

While logistics sprawl has occurred in some provinces in the Randstad region, warehouse development in peripheral areas has not taken place at the expense of dense areas across the region. The dynamic expansion that we have described corresponds to logistics facilities’ decentralization but it has not led to logistics facilities in peripheral areas dominating those in dense urban centers. In Amsterdam and Rotterdam metro areas logistics activities relocated around urban centers. This situation continued to apply between 2007 and 2013. Despite the presence of a strong sprawl dynamic affecting warehouses in the provinces of Flevoland (The Hague) and Utrecht, logistics is overall well integrated within dense areas. We can assume that The Hague and Utrecht have been experiencing logistics development later than Amsterdam and Rotterdam, and that logistics sprawl there is related to suburban sprawl. Overall, the Randstad example shows that dense areas can remain attractive for logistics activities. Analysis of the spatial dynamics in the Randstad reveals above all the existence of a form of urban logistics in which logistics activities remain a permanent feature of dense areas and in which areas of medium density undergo development. This can be called "logistics intensification".

Two forms, at least, of logistics urbanization coexist in metropolitan areas in Europe: logistics suburbanization and logistics intensification. In the case of the Paris region, logistics sprawl leads to logistics suburbanization, meanwhile the Randstad region experienced logistics contraction and densification. Public policies can explain some of the differences between logistics suburbanization and logistics intensification.

5 Conclusion

It has been observed in the literature that in many of the world's cities the development of logistics facilities has taken place in the outskirts as part of a process of logistics sprawl. However, it is noteworthy that most of these analyses were made in the case of monocentric urban regions, which contain one central urban area, highly developed suburbs, and less developed exurban areas. By analyzing changes in the location of warehouses in the polycentric Randstad region between 2007 and 2013, we highlighted the existence of another "model". Logistics intensification has occurred in the main urban areas (Rotterdam and Amsterdam) of the Randstad, with different distribution densities and center/suburb relationships. We have shown that each metro area of the Randstad has its own dynamic in terms of logistics. In the Randstad, logistics activities and warehouses tend to concentrate in (close to) metro areas, and we may relate that to its polycentric structure as well as (under links that remain to be specifically analyzed) its situation as a European freight hub and gateway. The specific tradition of land use control in the Randstad may also explain the intensification of logistics development in or close to the urban areas.

The different forms of logistic urbanization reveal a variable level of warehouse integration in metropolitan structures. Logistics activities are not condemned to "flee" into the suburbs. Under certain conditions, these activities can stay in the denser parts of a conurbation. We therefore formulate as a general conclusion that the diversity in the forms taken by logistics development is not simply due to the logistics and freight transport system, but also depends on the intrinsic characteristics of the regional spatial structure as well as local planning and land use policies.

This research cannot provide a final conclusion on the different types of logistics spatial developments according to the monocentric or polycentric nature of an urban region. In our two cases, the Paris region and the Randstad region, the intensification of logistics reflects a polycentric urban region while logistics suburbanization appears to be a symptom of a monocentric metropolitan area, but other factors, as we have shown, have played a role. This work is a starting point. What remains to be done, and could be relevant for a subsequent research about polycentric urban regions, is to look at freight and logistics interdependencies (freight flows to and from warehouses, main freight generators and consumption areas) between the four different provinces of the Randstad.

Acknowledgement

This article presents results of a research carried out for the MetroFreight Center of Excellence on urban freight. MetroFreight is financed by the Volvo Research and Educational Foundations (VREF). For this research, the author benefited from an academic exchange with TU Delft (September 2014).

References

Aljohania K, Thompson R (2016) Impacts of logistics sprawl on the urban environment and logistics: Taxonomy and review of literature. *Journal of Transport Geography* 57: 255–263. CrossRef.

Allain R (2004) Morphologie urbaine, géographie, aménagement et architecture de la ville. Armand Colin, Collection U, Paris

Allen J, Browne M, Cherrett T (2012) Investigating relationships between roadfreight-transport, facility location, logistics management and urban form. *Journal of Transport Geography* 24: 45–57. CrossRef.

Berroir S, Mathian H, Saint-Julien T, Sanders M (2008) La mobilité dans la construction du polycentrisme métropolitain. In: Thiérault M, Des Rosiers F (eds), *Information Géographique et dynamiques Urbaines 1*. Hermes-Lavoisier, collection, IGAT, 31–57

Beyer A (1999) Morphologies et dynamiques territoriales des services de messagerie. doctoral thesis, Université Paris XII

Bogaerts A, Dieleman F, Dijst M, Geertman S (2007) Economic deconcentration in a rational planning system: The Dutch case. In: Razin E, Dijst M, Vázquez C (eds), *Employment Deconcentration in European Metropolitan Areas: Market Forces versus Planning Regulations*. Springer, Dordrecht. CrossRef.

Bowen J (2008) Moving places: The geography of warehousing in the US. *Journal of Transport Geography* 16[6]: 379–87. CrossRef.

Burke G (1966) *Green Heart Metropolis: Planning the Western Netherlands*. Macmillan, Toronto

Champion A (2001) A changing demographic regime and evolving polycentric urban regions: Consequences for the size, composition and distribution of city populations. *Urban Studies* 38[4]: 657–677. CrossRef.

Cidell J (2010) Concentration and decentralization: The new geography of freight distribution in US metropolitan areas. *Journal of Transport Geography* 18[3]: 363–71. CrossRef.

Cidell J (2011) Distribution centers among rooftops: the global logistics network meets the suburban spatial imaginary. *International Journal of Urban and Regional Research* 35[4]: 832–851. CrossRef.

Dablanc L, Raimbault N (2015) Penser autrement la métropole logistique: questions d'aménagement et d'urbanisme. In: Dablanc L, Frémont A (eds), *La métropole logistique, Le transport de marchandises et le territoire des grandes villes*. Armand Colin, Paris

Dablanc L, Rakotonarivo D (2010) The impacts of logistic sprawl: How does the location of parcel transport terminals affect the energy efficiency of goods' movements in Paris and what can we do about it? *Procedia. Social and Behavioral Sciences* 2[3]: 6087–6096. CrossRef.

Dablanc L, Ross C (2012) Atlanta: A mega logistics center in the Piedmont Atlantic Megaregion (PAM). *Journal of Transport Geography* 24: 432–442. CrossRef.

Davoudi S (2003) Polycentricity in European spatial planning: from an analytical tool to a normative agenda. *European Planning Studies* 6[4]: 365–377. CrossRef.

Davydenko I, Tavasszy L, Quak H (2013) A quantitative model for exploration of logistics sprawl of the future. Metrans conference http://www.metrans.org/sites/default/files/-Metrans_Davydenko_Tavasszy_Quak.pdf

DRIEA (2009) Étude sur l'estimation et la caractérisation du parc d'entrepôts en Ile-de-France. http://www.driea.ile-de-france.developpement-durable.gouv.fr/IMG/pdf/-Estimation_parcentrepots_GRECAM_cle651884.pdf (last accessed Feb 17, 2017)

Frémont A (2015) Introduction. In: Dablanc L, Frémont A (eds), *La métropole logistique. Le transport de marchandises et le territoire des grandes villes*. Armand Colin, Paris, 9–18

Hall P (2008) How polycentric are mega-city regions? Nova terra, January, 24-29

Hall P, Hesse M (2013) Cities, flows and scale. policy responses to the dynamics of integration and disintegration. In: Hall P, Hesse M (eds), *Cities, Regions and Flows*. Routledge, London, 247–259

Heitz A, Dablanc L (2015) Logistics spatial patterns in Paris: The rise of the Paris basin as a logistics megaregion. *Transportation Research Record: Journal of the Transportation Research Board* 2477: 76–84

Hesse M (2002) Location matters. ACCESS magazine. University of California Transportation Center, 21, 22-26

Hesse M (2004) Land for logistics: locational dynamics, real estate markets and political-regulation of regional distribution complexes. *Tijdschrift voor economische en sociale geografie* 95[2]: 162–173. CrossRef.

Kühn M (2003) Greenbelt and green heart: separating and integrating landscapes in European city regions. *Landscape and urban planning* 64: 19–27. CrossRef.

Lasserre F (2004) Logistics and the internet: transportation and location issues are crucial in the logistics chain. *Journal of Transport Geography* 12[1]: 73–84. CrossRef.

Le Nechet F, Aguiléra A (2012) Forme urbaine et mobilité domicile-travail dans 13 aires urbaines françaises: une analyse multiéchelle. *Recherche Transports et Sécurité* 28[3-4]: 259–270

Merenne-Schoumaker B (2008) La localisation des grandes zones de logistique. *Bulletin de la Société géographique de Liège* 49: 31–40

Ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieu (2008) Randstad Holland in 2040, environment and spatial planning. Strategic document

Mitchell A (2005) Manuel ESRI d'analyse SIG. 2. Paris, ESRI Press

Nicot B (1996) Une mesure de l'étalement urbain en France, 1982-1990. *Revue d'Economie Régionale et Urbaine* 1: 71–98

OECD (2007) Randstad Holland, Netherlands. OECD Territorial Reviews

Priemus H, Visser J (1995) Infrastructure policy in the Randstad Holland: struggle between accessibility and sustainability. *Political Geography* 14[4]: 363–377. CrossRef.

Raimbault N (2015) Grande distribution: entre performance logistique et contrainte foncière. In: Dablanc L, Frémont A (eds), *La métropole logistique, Le transport de marchandises et le territoire des grandes villes*. Armand Colin, 161–178

Raimbault N, Bahoken F (2014) Quelles places pour les activités logistiques dans la métropole parisienne? *Territoire en Mouvement* 23-24: 53–74. CrossRef.

Rijksdienst voor het Nationale Plan (1958) Netherlands. werkcommissie westen des lands. Staatsdrukkerij en uitgeverijbedrijf

SDRIF – Schéma directeur de la région Île-de-France (2013) Schéma directeur de la région Île-de-France. https://www.iledefrance.fr/action-quotidienne/construire-ile-france-2030-0 (last accessed March 12, 2017)

van den Heuvel FP, de Langen PW, van Donselaar KH, Fransoo JC (2013) Spatial concentration and location dynamics in logistics: the case of a Dutch province. *Journal of Transport Geography* 28: 39–48. CrossRef.

van der Burg A, Vink B (2008) Randstad Holland 2040. 44th ISOCARP congress

Ville de Paris (2016) Plan Local d'Urbanisme de Paris. Document stratégique de planification

Woudsma C, Jakubicek P, Dablanc L (2015) Logistics sprawl in North America: methodological issues and a case study in Toronto. *Transportation Research Procedia* 12: 474–488. CrossRef.

Woudsma C, Jensen J, Kanaroglou P, Maoh H (2008) Logistics land use and the city: a spatial-temporal modeling approach. *Transportation Research Part E, Logistics and Transportation Review* 44[2]: 277–297. CrossRef.

REGION
The Journal of ERSA
Powered by WU

Volume 4, Number 1, 2017, 109–127
DOI: 10.18335/region.v4i1.107
journal homepage: region.ersa.org

Revisiting the Boston data set – Changing the units of observation affects estimated willingness to pay for clean air

Roger Bivand[1]

[1] Norwegian School of Economics, Bergen, Norway, (email: roger.bivand@nhh.no); Adam Mickiewicz University, Poznań, Poland

Received: 2 December 2015/Accepted: 6 April 2017

Abstract. Harrison, Rubinfeld (1978c) used a hedonic model to find out how house values were affected by air pollution in Boston, when other variables were taken into consideration. Their primary interest was in estimating willingness to pay for cleaner air. They chose to use 506 census tracts as units of observation because median house values for these units of aggregation were published in the 1970 census tabulations. Air pollution values from the model output, represented by nitrogen oxides (NOX), were available for 122 model output zones, of which only 96 fell within the study area defined by the chosen census tracts. These NOX values were then assigned proportionally to all census tracts falling within each model output zone. By re-aggregating the house value data to the 96 air pollution model output zones and re-fitting the regression model, the total impact of air pollution on house values, and thus the estimated willingness to pay, increases markedly. By extending the analysis to include spatially lagged independent variables, the total impact of air pollution on median house values, and consequently on the willingness to pay analysis, increases by over three times. Use of weighting to adjust the units of observation for the relative numbers of housing units behind each median house value further buttresses this conclusion. It is shown conclusively that the choice of observational units matters crucially for the estimation of economic parameters of interest in this data set.

1 Introduction

The Harrison, Rubinfeld (1978c) Boston housing data set has been widely used because of its availability from Belsley et al. (1980) and Statlib, and was further discussed by Pace, Gilley (1997) and Gilley, Pace (1996). The original paper has been highly cited, with almost 400 journal references, mostly by data analysts using the original data set for testing estimation methods, rather than by environmental or urban economists. The underlying research question was the estimation of willingness to pay (WTP) for clean air, using air pollution levels and house values in a hedonic regression (see also Harrison, Rubinfeld 1978b,a, which are not highly cited). Differences in house values for otherwise similar observations but differing pollution levels should permit the estimation of a coefficient for the calibration of WTP for clean air. As Pace, Gilley (1997, p. 337) showed clearly, the air pollution coefficient estimate in their fitted models changed when residual spatial autocorrelation was taken into account (from -0.0060 to -0.0037), as did its standard error (from 0.0012 to 0.0016).

Repeated attempts to fit spatial econometrics models to the 506 observations of census tracts in the original data set, and to take account of censoring as noted by Gilley, Pace (1996), have not resolved the question of what WTP would be in a model not subject to serious mis-specification issues. The modelled data set suffers from spatial autocorrelation in many ways, as well as heteroskedasticity, and there are further questions about functional form.

The use of spatial econometrics models described by Halleck Vega, Elhorst (2015) and LeSage (2014) including the spatially lagged dependent variable, spatially lagged errors, and spatially lagged independent variables does not seem to be successful. The best fit is achieved with the General Nested Model (GNM), with the total impact reaching -0.0093, somewhat larger than the least squares coefficient value (-0.0066), but much less significant. The levels of autocorrelation seen in the air pollution variable are very high, something which could have a reasonable basis in the data generation process of the variable, but which deserves investigation. Attempting to instrument for the possible errors-in-variable status of the air pollution variable, as suggested by Anselin, Lozano-Gracia (2008) using methods provided by Drukker et al. (2013) and Piras (2010) also results in unsatisfactory outcomes: the air pollution total impact is insignificant and positive (0.0042) in a GMM SARAR model with air pollution taken as subject to errors-in-variables and instrumented by a quadratic trend surface.

The levels of spatial autocorrelation present in the GNM are very large and highly significant ($\rho_{\text{Lag}} = 0.7436$, $\rho_{\text{Err}} = -0.2842$),[1] but do not have any clear economic interpretation. Dropping the spatially lagged independent variables from the model yields equally significant spatial coefficients, but with a sign flip ($\rho_{\text{Lag}} = 0.2808$, $\rho_{\text{Err}} = 0.4889$). These results mirror those found in the GMM SARAR with the air pollution subject to errors-in-variables, with spatial coefficients large in absolute value and highly significant ($\rho_{\text{Lag}} = 0.3867$, $\rho_{\text{Err}} = 0.3645$).

In revisiting the Harrison, Rubinfeld (1978c) Boston housing data set in this study, an attempt will be made to establish the reason for the presence of this very strong spatial autocorrelation. Gibbons, Overman (2012) argue that spatial econometrics is too often applied without sufficient consideration of the underlying economics; in this paper, we will rather consider the configuration of the units of observation. Is the strength of spatial autocorrelation observed in this data set a feature of the census tract observations themselves, or has it been introduced or strengthened by changes in the observational units used for the different variables? Our focus will be on the choices of observational entities made in assembling the original data set, and on an alternative that arguably should be more relevant for the data generation process of the air pollution variable. Having re-established an approximation to the model output zones from which the air pollution variable levels were taken, it will be shown that much of the puzzling spatial autocorrelation is removed. A further question to be considered is whether weights should be used to account for the very different numbers of housing units found in each observational entity, taking up the challenge given by Solon et al. (2015).

2 Observation entities in Harrison, Rubinfeld (1978c)

In order to approach WTP for cleaner air, Harrison, Rubinfeld (1978c) used a hedonic regression including air pollution levels with house values as the dependent variable. They use a data set for most of the Boston SMSA in 1970 at the census tract level of aggregation. The data were made available by Belsley et al. (1980, pp. 229–261) in the form in which they appear to have been analysed. Pace, Gilley (1997) and Gilley, Pace (1996) found that there were errors in Belsley et al. (1980) and the Statlib data file, and that the house value data were censored.

[1] ρ_{Err} is the spatial autoregressive coefficient on the error term, ρ_{Lag} is the spatial autoregressive coefficient on the spatially lagged dependent variable

Table 1: Five-number summary of counts of owner-occupied one-family housing units by observational unit; observational units 506 Census tracts and 96 approximate TASSIM zones.

	minimum	lower quartile	median	upper quartile	maximum
506 Census tracts	5.00	115.00	511.50	1155.00	3031.00
96 TASSIM zones	25.00	1624.00	2926.50	5189.50	12411.00

2.1 House values

Harrison, Rubinfeld (1978c) used median house values in 1970 USD for 506 census tracts in the Boston SMSA for owner-occupied one-family houses; census tracts with no reported owner-occupied one-family housing units were excluded from the data set. Here the values are not at the micro-level but medians from census tracts from the 1970 US Census (for "owner-occupied one-family housing"). The relevant question is H11: "If you live in a one family house which you own or are buying — What is the value of this property? That is, how much do you think this property (house and lot) would sell for if it were for sale?"[2]. H11 was answered by crossing off one grouped value alternative, ranging from under \$5,000 to over \$50,000. Tracts with weighted median values in these upper and lower alternative value classes are censored. The house value data have census tract support, and are median values calculated from group counts from the alternatives offered in H11.

The published census tract tabulations show the link between question H11 and the Statlib-based data (after correction).[3] The median values tabulated in the census report can be reconstructed from the tallies shown in the same Census tables fairly accurately using the `weightedMedian` function in the **matrixStats** in R, using linear interpolation, and midpoint values of \$3,500 and \$60,000 for the left- and right-censored intervals.[4]

Two tracts are entered as having a median house value below \$5,000, and 15 have median values over \$50,000, as was pointed out by Gilley, Pace (1996). One tract has a median of exactly \$50,000, with 31 houses below the right-censored boundary, and 31 above. Having access to the Census value group counts by tract means that alternative aggregations of house value — the dependent variable in the analysis — may be constructed using the underlying data.

The effectiveness of the study was prejudiced by the fact that areas of central Boston with the highest levels of air pollution also lose house value data, either because of tract exclusion (no one-family housing units reported) or right or left censored tracts. Figure A.2 shows the impact of censoring in central Boston, where the highest air pollution values were predicted, with few tracts represented by weighted median house values. The excluded tracts contained no owner-occupied one-family housing units, or too few to tabulate, but may have included rented housing, which was not considered in the original study.

A further point made by Harrison, Rubinfeld (1978c, pp. 88, 98–101), is that the number of housing units by tract varies greatly. They tried using weighted regression, using the logarithm of the counts of one-family houses by tract as the weighting variable, and noted some change in coefficient values. This step was taken to attempt to check the results for robustness to heteroskedasticity. Figure A.5 and Table 1 show the considerable disproportions present in the data, both for the original choice of census tracts as observational units, and for the alternative Transportation and Air Shed SImulation Model (TASSIM) zones proposed below.

The data used by Harrison, Rubinfeld (1978c) took the census tract as the observational

[2]Question reproduced in Figure A.1; https://www.census.gov/history/pdf/1970_questionnaire.pdf, p. 5.

[3]Census of Population and Housing-1970-Census Tracts: Part 3 Binghamton, N.Y.-PA.SMSA-Cedar Rapids, Iowa SMSA, http://www2.census.gov/prod2/decennial/documents/39204513p3_TOC.pdf, Sections 5 and 6, PHC(1)-29, table H1.

[4]http://www2.census.gov/prod2/decennial/documents/00116813p1.zip, Chapter 5, Census Users' Dictionary, p. 118.

unit as we have seen. The US Census is organized using a nested system of blocks, block groups and tracts. Data for house value counts for blocks are often not available because the numbers in each value group are too small for publication. Table A.1 shows the relative distribution of counts of blocks and housing units per block based on data extracted from the National Historical Geographic Information System.[5] The choice of census tracts by the original authors as the smallest feasible areal unit seems justified, especially as most of the other variables they used were not available at the block level, especially air pollution.

2.2 Air pollution

The data on air pollution concentrations were obtained from the Transportation and Air Shed SIMulation model (TASSIM) applied to the Boston air shed (Ingram, Fauth 1974). That study was conducted to simulate the possible consequences of abatement policies affecting road traffic. It used data on vehicle and point-source emissions combined with meteorological data to generate a number of mean air pollution concentration surfaces, which were then calibrated to values from monitoring stations. The calibrated model results were obtained for 122 zones, and assigned proportionally to the 506 census tracts. Values were taken for particulate matter (PART) and nitrogen oxides (NOX), and analysis proceeded using only NOX. The NOX values in the published data sets are in units of 10 ppm (10 parts per million), and were then multiplied by 10 again in the regression models to yield parts per 100 million (pphm).[6]

Many of the smaller tracts belong to the same TASSIM zones; this is a clear case of change of support, with possibly different spatial statistical properties under the two different entitation schemes (Gotway, Young 2002). Harrison, Rubinfeld (1978c, p. 86, footnote 14) do comment that "... the true correlation between NOX and PART is somewhat overstated because the TASSIM model generates data for 122 zones, not 506 census tracts. Translating zonal data into census tracts tends to overstate the correlation because relatively more census tracts are located in center city zones in which PART and NOX levels tend to be most highly correlated." This is not directly related to the modifiable areal unit problem (Wong 2010); the relationship between this case and the modifiable areal unit problem will be discussed in the conclusion below.

Since the data set from Belsley et al. (1980) does not include PART, nor the identifiers of the TASSIM zones underlying the assignation of copied values to census tracts, it is not possible from the data as they stand to retrieve the zones with full certainty. We can, however, aggregate contiguous census tracts with identical values of NOX, giving 96 approximated TASSIM zones, for which we can aggregate grouped house value counts, and calculate median values using the same procedure as that used at the census tract level of resolution.

A two-part report[7] details the use of the TASSIM simulation model (Ingram, Fauth 1974, Ingram et al. 1974). Both of these volumes include line-printer maps of the TASSIM zones, and the Fortran code in volume 2 (Ingram et al. 1974, pp. 183–185) shows the links between the 122 TASSIM zones and the line printer output. There is no description of the projection used, so relating these polygons to tract boundaries is not exact. Western TASSIM zones appear to lie outside the Boston SMSA tracts included in the 506 census tract data set. An affine transformation between ground control points in the map of 506 census tracts, and guessed equivalents in the line printer TASSIM zones is shown in Figure A.3. The remaining discrepancies appear to come from merging contiguous census tract data set entities with the same NOX values which actually belonged to more than one TASSIM zone, and from poor matching because of difficulties in locating ground

[5]Minnesota Population Center. National Historical Geographic Information System: Version 2.0. Minneapolis, MN: University of Minnesota 2011.

[6]US environmental agencies today typically use ppb (parts per billion), while international and European agencies typically use μg/m^3, with 40μg/m^3 the current WHO annual average recommended threshold (about 19.5 ppb under given assumptions about air pressure and temperature). Air pollution in Boston has declined by over half since the 1970's, but in our data set, all the TASSIM zones had values between two and four times today's international threshold for average annual NOX exposure.

[7]Thanks to Ewa Wołynska, Central Connecticut State University Library, for locating and scanning this report.

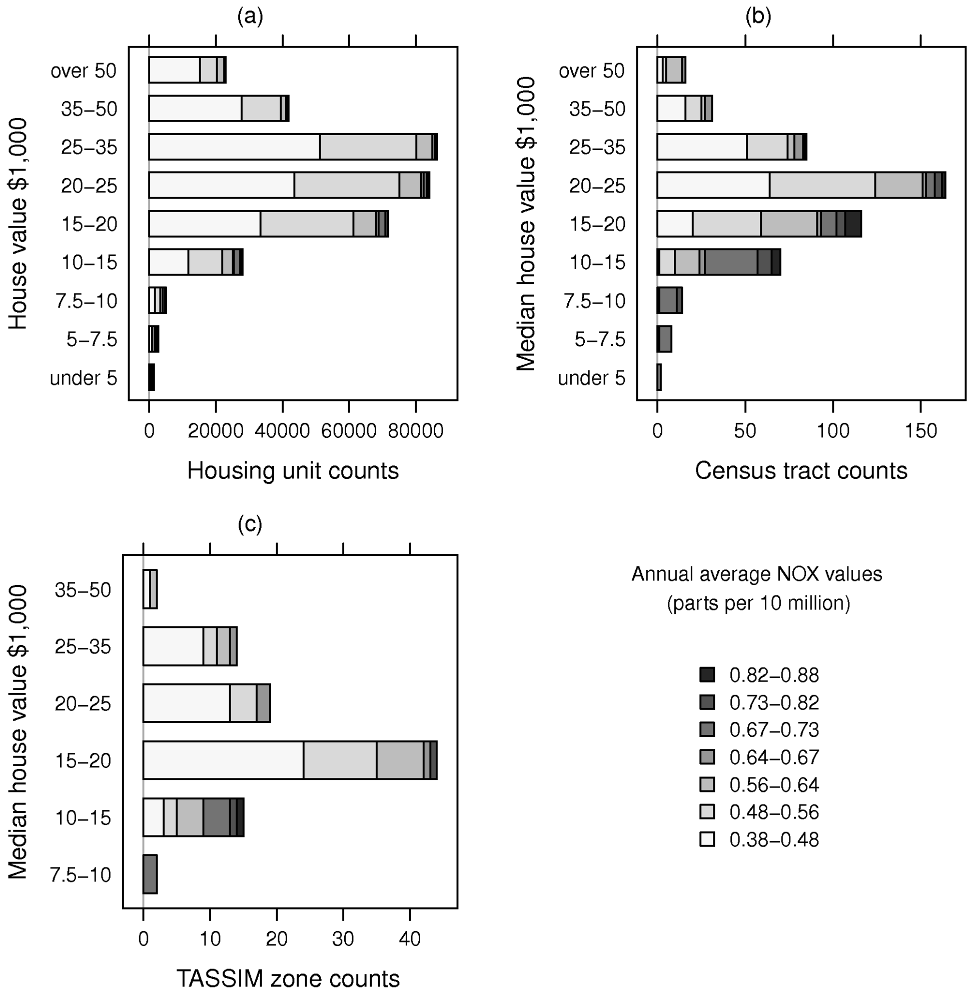


Figure 1: Bar charts by H11 classes (under 5, 5–7.5, 7.5–10, 10–15, 15–20, 20–25, 25–35, 35–50, over$50 thousand 1970) and seven natural breaks average annual NOX classes given in parts per 10 million (also used for fill colors) — (a): counts of housing units; (b): counts of census tracts by median house values; (c): counts of TASSIM zones by median house values.

control points on the line printer output map; there are few discrepancies remaining, less than ten percent of the TASSIM zone entities.

Figure 1 uses grey shades to visualize the possible impact of using weighted regression, with panel (a) showing — with the chosen class intervals for NOX — the actual relationship between house value and NOX. The NOX levels and house value class counts are taken from the census tract data set. In the remaining panels, the bar lengths are proportional to the counts of spatial entities with median house values falling into the H11 house value classes. NOX fill shades represent levels of air pollution. Tracts and TASSIM zones with higher levels of air pollution typically have many fewer housing units. It can be seen that the numbers of housing units with higher levels of air pollution are proportionally less in (a) than in (b) or (c), and that the largest value class for housing units was $25,000–35,000, for $20,000–25,000 for census tracts, and $15,000–20,000 for approximate TASSIM zones. An argument for weighing the hedonic regressions is to shift the interpretative basis back to that of panel (a), that is the actual numbers of housing units subject to different levels of air pollution.

Figure 2 shows clearly that the study of the relationship between NOX and house value will be impacted by "copying out" NOX values to census tracts, as noted by Harrison, Rubinfeld (1978c, p. 86, footnote 14) and mentioned above. Even if we were to use more

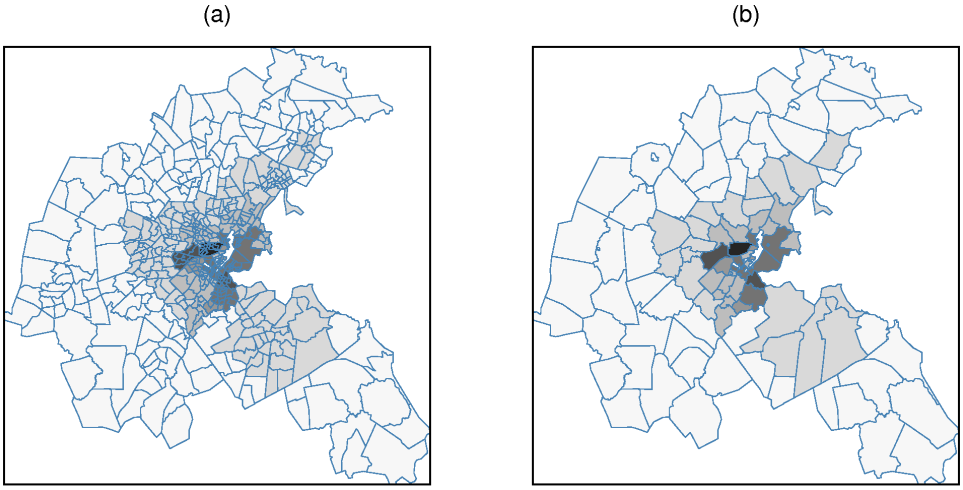


Figure 2: NOX values in parts per 10 million shown using the same class intervals and color fill as Figure 1 for two sets of observation units: (a): 506 census tracts and (b): 96 approximate TASSIM zones.

class intervals in these choropleth maps, the visual impression would be the same, because the underlying data have support approximated by the 96 TASSIM zones, not by the 506 census tracts.

Figure A.4 shows changes in the visual impression given by aggregating the H11 median house value counts to approximate TASSIM zones before calculating interpolated weighted median values. It also shows the censored census tracts for which we have no reliable median values, as the values taken depend on the assumed under/over interval midpoints. Once we aggregate to TASSIM zones, there are no longer any censored median values, because the weighted medians no longer fall outside the bounded range between $5,000–50,000.

2.3 Other independent variables

Besides NOX, the other census variables included in the hedonic regression to account for median house values are the average number of rooms per house, the proportion of houses older than 1940, the proportion low-status inhabitants in each tract, and the Black proportion of population in the tract — originally expressed as a broken-stick relationship, but here taken as a percentage. The crime rate is said to be taken from FBI data by town, but which is found on inspection to vary by tract. The distance from tract to employment centers is derived from other sources, as is the dummy variable for tracts bordering Charles River.

Other independent variables are defined by town, with some also being fixed for all towns in Boston. The town aggregates of census tracts are used in many of the census report tabulations, and of the 92 towns, 17 only contain one census tract, while one town contains thirty census tracts. The variables are the proportion of residential lots zoned over 25,000 sq. ft, the proportion of non-retail business acres, accessibility to radial highways, full-value property-tax rate per $10,000, and pupil-teacher ratio by town school district. These variables are also "copied out" to tracts within towns, but do not coincide completely with the approximate TASSIM zones. In the case of 80 approximate TASSIM zones aggregated from census tracts, the boundaries do coincide exactly with town boundaries. However, for the remaining 12 towns and 16 TASSIM zones, there are overlaps between more than one town and TASSIM zone, mostly in Boston itself. The exact match between town boundaries defined using census tracts, and approximated TASSIM zones also constructed using census tracts is not necessarily an indication that towns were used as TASSIM zones, but as we cannot reconstruct the actual model output zones exactly, we assume that the difference is without importance for this study. Using

TASSIM zones for analysis should therefore also reduce the levels of autocorrelation induced by "copying out" town values to tracts within towns.

Table A.2 shows the descriptive statistics for the variables used in the 489 observation census tract data set omitting tracts with censored median house values. These covariates were aggregated to approximate TASSIM zones using weighted averages, where the weights are the tract population counts. The Charles River dummy was aggregated by taking the maximum value of any tract included in the approximate TASSIM zone. It would be possible to punch[8] more census data for some of the variables, but not all the variables used are present in the census tables available online. Table A.3 shows the descriptive statistics for the variables used in the 96 observation approximate TASSIM zone data set.

3 Applying weighted spatial econometrics models to the Boston data sets

Pace, Gilley (1997) felt that it should be worthwhile to check whether the original model was spatially misspecified. They considered that the use of spatial aggregate units as observations might involve spillovers of some kind, chiefly in the house values used – neighboring census tracts may have similar values for a number of reasons. Had the included explanatory variables accounted for the similarities between neighbors, there might not have been any reason to go further, but the residuals were found to be spatially highly patterned. So now we will turn to spatial econometrics methods to try to unravel the question of the "real" link between house values and NOX. We use row-standardized contiguity neighbors derived from the map of census tracts, omitting the censored tracts which leads to one tract having no neighbors, and from the map of merged census tracts constituting approximate TASSIM zones.

3.1 Spatial econometrics models with case weights

The models chosen here are unweighted and weighted variants of least squares (OLS) and the spatial error model (SEM), both also extended by including spatially lagged independent variables (SLX and SDEM respectively). The arguments for and against the use of weights have been reviewed recently by Solon et al. (2015). They distinguish between situations in which the analyst is most interested in estimating descriptive statistics for a population, and the estimation of causal effects, for example through correcting for heteroskedasticity. It may be argued that the hedonic model used by Harrison, Rubinfeld (1978c) is less concerned with establishing the causal effect of air pollution on house values than with estimating the coefficient expressing how house values would respond to changes in air pollution, thus providing insight into WTP for clean air. Given the very large variations in counts of housing unit valuations underlying each median value demonstrated above, it seems reasonable to consider the numbers of housing units by observation as case weights (it was noted above that Harrison, Rubinfeld (1978c, pp. 88, 98–101) checked logarithms of these counts as weights in a robustness test but found little difference from OLS).

There are two reasons for choosing not to include the spatially lagged median house value dependent variable in the models considered. The first is based on LeSage (2014), and the probability that the aggregate nature of the dependent variable makes it seem more reasonable to consider local spillover specifications. The "copying out" of covariates across multiple tracts from the different entitation schemes for TASSIM output zones and for towns can arguably be seen as being better represented by local rather than global spillovers. The second reason is pragmatic, that weighted spatial regression code in the **spdep** package in R is so far only implemented for the spatial error (SEM) and by extension the spatial Durbin error model (SDEM).[9]

The dependent variable representing house value is taken as the logarithm of median house values in 1970 USD, and the air pollution variable is the square of NOX in parts

[8] Data entry in 1970 was by punch card machines; the term means to enter data into a computer system.

[9] The `spautolm` function was originally written to replicate results in Waller, Gotway (2004, p. 378), and weights have now been added to the `errorsarlm` to match (relative differences in coefficient estimates are 1e-8, because of different underlying numerical methods).

Table 2: Likelihood ratio test p-values for nested pairs of fitted models and two weighing choices for the 96 TASSIM zones and the 489 census tract data sets; model pairs are shown as NESTED:NESTING combinations, and the test is for the lack of significance in fit between the nested and nesting models.

	TASSIM	TASSIM weighted	Census	Census weighted
OLS:SLX	4.685e-09	5.246e-05	< 2.2e-16	1.045e-11
OLS:SEM	0.06099	0.0009402	< 2.2e-16	< 2.2e-16
OLS:SDEM	1.062e-08	9.246e-05	< 2.2e-16	< 2.2e-16
SLX:SDEM	0.789	0.6512	< 2.2e-16	< 2.2e-16
SEM:SDEM	1.966e-08	0.002528	1.388e-10	6.28e-05

per hundred million; these and all other variables are represented as in the original study (with the exception of the Black proportion of population). We will now present briefly the models used. Assuming that the variance of the disturbance term is constant, we start from the standard linear regression model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim N(0, \Omega), \Omega = \sigma^2\mathbf{I}$$

where $\mathbf{y}$ is an $(N \times 1)$ vector of observations on a dependent variable taken at each of N locations, $\mathbf{X}$ is an $(N \times k)$ matrix of exogenous variables, $\boldsymbol{\beta}$ is an $(k \times 1)$ vector of parameters, and $\boldsymbol{\varepsilon}$ is an $(N \times 1)$ vector of disturbances. The spatial error model (SEM) may be written as (Ord 1975):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \qquad \mathbf{u} = \rho_{\mathrm{Err}}\mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon},$$

where $\mathbf{y}$ is an $(N \times 1)$ vector of observations on a dependent variable taken at each of N locations, $\mathbf{X}$ is an $(N \times k)$ matrix of exogenous variables, $\boldsymbol{\beta}$ is an $(k \times 1)$ vector of parameters, $\boldsymbol{\varepsilon}$ is an $(N \times 1)$ vector of disturbances and ρ_{Err} is a scalar spatial error parameter, and $\mathbf{u}$ is a spatially autocorrelated disturbance vector with constant variance and covariance terms specified by a fixed $(N \times N)$ spatial weights matrix $\mathbf{W}$ and a single coefficient ρ_{Err}:

$$\mathbf{u} \sim N(0, (\mathbf{I} - \rho_{\mathrm{Err}}\mathbf{W})^{-1}\Omega(\mathbf{I} - \rho_{\mathrm{Err}}\mathbf{W}^{\top})^{-1}), \Omega = \sigma^2\mathbf{I}$$

In both cases, the Durbin forms (SLX, SDEM) are defined by augmenting the matrix of independent variables $\mathbf{X}$ with its spatial lag $\mathbf{WX}$, now using $[\mathbf{X}, \mathbf{WX}]$ instead of just $\mathbf{X}$ (the intercept is included only once in the case of row-standardized spatial weights). The weighted versions are formed by altering $\Omega = \sigma^2\mathbf{I}$ by replacing the identity matrix by a diagonal matrix of the inverses of known case weights reflecting the relative "size" of the observations (Waller, Gotway 2004). The variance term σ^2 is still estimated, but with the assumption of uniform variance replaced by variance proportional to the inverse of known case weights.

3.2 Model diagnostics

Comparison of models fitted using different models to different aggregations of the same data may be handled qualitatively, as there does not seem to be an agreed basis in statistics or econometrics from which to proceed. The Durbin versions of OLS and SEM will subsequently be termed SLX and SDEM (LeSage 2014, Halleck Vega, Elhorst 2015). The weighted and unweighted OLS and SLX models were fitted by least squares, and the weighted and unweighted SEM and SDEM models were fitted by maximum likelihood. The SDEM model nests the others: OLS, SLX and SEM, but SEM only nests OLS, not SLX. The values of likelihood ratio tests are shown in Table 2 by data set and the use of weighting. In the 489 census tract data set, the SDEM model appears to be dominant in both the weighted and unweighted cases, while the SLX model appears to dominate in the 96 TASSIM zones data set.

More complete tables of diagnostic tests for groups of models are presented in the Appendix below, but discussed briefly here. Tables A.4 and A.5 show the AIC values for fitted models for two sets of entities: 489 census tracts and 96 approximate TASSIM zones, and approximately the same data. These values, although also based on the log likelihood, penalize models with more independent variables, here the SLX and SDEM models including the spatially lagged independent variables. The data for median house values are constructed in exactly the same way, as is the air pollution variable, while other variables for TASSIM zones are population-weighted aggregates of the values in the census tract data set. The best model fit may be held to be indicated by the lowest AIC value among comparable models.

In Table A.5, the spatial error models no longer outperform the models without a spatially lagged error term, and the spatial coefficients of the spatial error models are not significant (see Table A.10). The SLX models do outperform their non-weighted counterparts, subject to the remark above about the comparability of these models. If we prefer the census tract data set, we might conclude that the weighted SDEM model is to be preferred, but if we choose the approximate TASSIM zones, our choice would be the weighted SLX model, including the spatially lagged independent variables.

Table A.6 shows Breuch-Pagan tests for heteroskedasticity for models fitted by least squares; columns 3 and 4 test against housing unit counts per aggregate entity as the source of heteroskedasticity. The OLS and SLX fitted models for the 489 census tracts show strong heteroskedasticity, which is reduced but not removed when modelled by housing unit counts. For the two models fitted using the 96 TASSIM zones, heteroskedasticity is only present for OLS, but not for SLX, and weighing with housing unit counts largely removes heteroskedasticity.

Standard deviates of Moran's I test for spatial autocorrelation in least squares regression residuals for weighted and unweighted OLS and SLX models are shown in Table A.7. All the test results for the 489 census tract data set are highly significant, while only the 96 TASSIM zones weighted OLS model residuals show any significant residual autocorrelation using this test. Table A.8 adds to this for the unweighted models estimated using least squares, showing the results of robust Lagrange multiplier tests. For the 96 TASSIM zones data set, only the robust test for an omitted lag coefficient is borderline significant, while all the robust tests for the 489 census tract data set were highly significant apart from the robust test for an omitted spatial lag coefficient for the SLX model.

Using the Hausman test given by Pace, LeSage (2008), and extended for case weights, it is found that all the tests for the 489 census tract data set for spatial and spatial Durbin error models (SEM, SDEM), whether weighted or not, reject the null hypothesis of no differences between least squares and spatial model coefficients on independent variables. In the case of the 96 TASSIM zones data set, only the test result for the unweighted SEM model is significant (Table A.9). Finally, Table A.10 shows spatial coefficient ρ_{Err} values and Wald test p-values for spatial error and spatial Durbin error models. All the ρ_{Err} values are highly significant for the 489 census tract data set, but only the SEM (both weighted and unweighted) ρ_{Err} values are significant for the 96 TASSIM zones data set.

Based on these diagnostics, it is clear that all the models fitted for the 489 census tract data set have serious mis-specification issues, and that the models fitted for the 96 TASSIM zones data set appear to be more successful, especially those including the spatially lagged independent variables, whether weighted or not.

3.3 Interpretation of model results

While we cannot directly compare relative model mis-specification across the two sets of entities (census tracts and approximate TASSIM zones), we can compare coefficient values for the key variable of interest, air pollution, taken as the square of NOX in the original scaling (parts per hundred million, pphm). For brevity, and because our focus here is on the consequences of choices of units of observation for estimates of the air pollution coefficient, other results are not given here, but may be obtained from the reproduction code. Note that by reducing the number of observations by a factor of five, the standard errors of the TASSIM zone coefficients are made much larger by construction. Further, note that standard errors reported below for the SEM and SDEM models are maximum likelihood estimates corrected by multiplication by $N/(N-k)$ for comparison with OLS and SLX standard error estimates.

Panel (a) of Figure 3 shows the coefficient values and 95% standard error bars for eight models excluding spatially lagged independent variables. The result for the unweighted SEM model for the census tract data set is not dissimilar from that given by Pace, Gilley (1997). The effect of residual spatial autocorrelation on the standard errors (and indeed on the coefficient values) is shown by comparing the OLS and SEM results for the census tract data set, regardless of whether weights are used.

When we move our attention to the Durbin models, including the spatially lagged covariates, we begin to be able to discern the consequences of the choice of observational units for inference about the air pollution variable. Panel (b) of Figure 3 shows the direct impacts, which are the NOX coefficient values from eight models. In the case of the 489 census tract data set, the coefficient values are positive and insignificant. For the 96 approximate TASSIM zones, the values are negative, as expected, and significant. It is the choice of observation units that makes the greater difference, larger than the inclusion or not of a spatial process in the disturbances, and larger than the inclusion or not of case weights to treat heteroskedasticity.

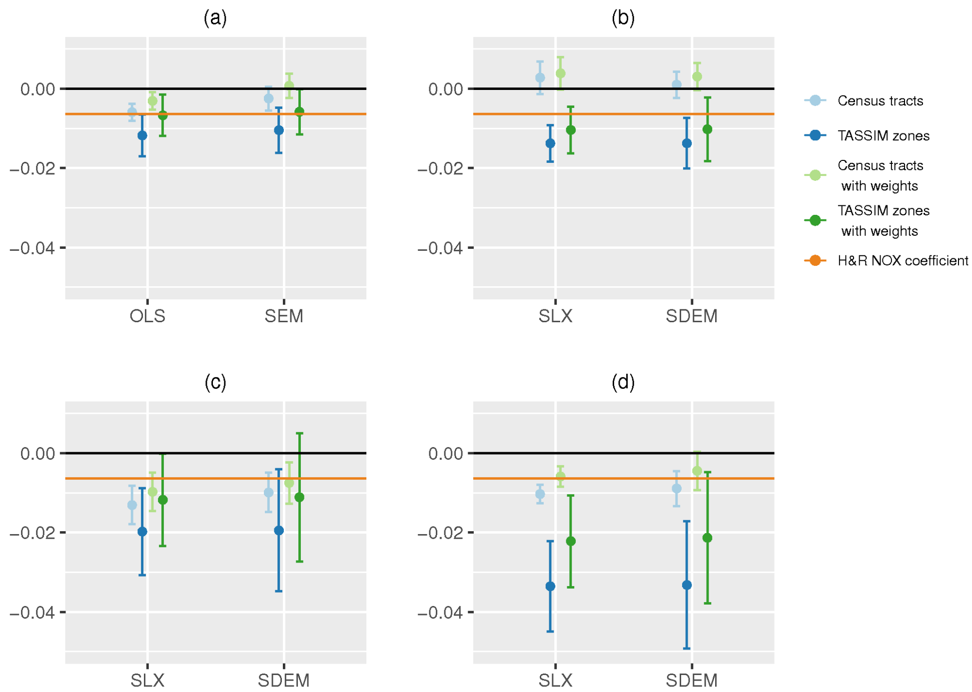


Figure 3: NOX coefficient values (a) and direct (b), indirect (c) and total (d) impacts and 95% standard error bars for OLS and SEM models, weighted and unweighted, for census tract and approximate TASSIM zone data sets; the Harrison, Rubinfeld (1978c, p. 100) coefficient value is shown in orange.

Panel (c) of Figure 3 shows the indirect impacts (the coefficients on the spatially lagged air pollution variable) and 95% times their standard errors. All are negative and the 489 census tract data set models are all significant. The standard errors of the 489 census tract NOX coefficients are much smaller than those of the models fitted using the 96 approximate TASSIM zones data set, not just because of the difference in numbers of observations. The NOX indirect impacts for the weighted models fitted using the 96 approximate TASSIM zones data set are at best marginally significant, so that with this data set and weighted regression, most of the "action" is in the direct impacts.

Finally, panel (d) of Figure 3 shows the total impacts for the eight models including the spatially lagged covariates, calculated using linear combination of the fitted model results for the NOX variable and its spatial lag. The total impact is simply the sum

Table 3: WTP for a one pphm reduction in NOX in 1970 USD (mean difference between base prediction and prediction with NOX reduced by 0.1 parts per ten million).

	OLS	SLX	SEM	SDEM
TASSIM	2118	6693	1866	6620
TASSIM, weighted	1157	4178	998	3999
Census tracts	1254	2223	520	1899
Census tracts, weighted	640	1240	-150	929

of the coefficient values, but the standard errors are calculated using the `estimable` function in the R **gmodels** package. The models fitted using the two different choices of observation entities differ considerably, with strong residual spatial autocorrelation in both the weighted and unweighted SLX models for the census tracts data set (see also TableA.10). The SDEM spatial autoregressive coefficients ρ_{Err} for the census tracts data set are 0.658 (standard error 0.0416) for the unweighted model and 0.62 (standard error 0.0444) for the weighted model. The equivalent values for the SDEM models for the approximate TASSIM zones data set are 0.0562 (standard error 0.155) for the unweighted model and 0.0964 (standard error 0.153) for the weighted model. The choice of observation entities is driving the value of the spatial error coefficient and inference on the appropriateness of its inclusion.

If we choose the approximate TASSIM zones data set, and drop the SDEM specification including ρ_{Err}, the spatial error coefficient, in favor of the SLX specification, we still need to choose whether to use the numbers of housing units as weights for the zones, or not to do so (implicitly upweighting zones with relatively fewer housing units, and downweighting those with many). Table A.5 shows that the AIC values differ. Although AIC values give some guidance, and are based on log likelihood values that take account of the given weights, the choice between the two models depends on the analyst's prior choice of weights. This suggests that Bayesian methods may well be relevant to permit better insight into this question in future research. If we take the 96 observation weighted SLX specification, the total impact of NOX is -0.02217 (standard error 0.005901), with equivalent values for the unweighted case: -0.03353 (standard error 0.005798). These values are substantially larger in absolute terms when compared with that found in Harrison, Rubinfeld (1978c, p. 100, coefficient -0.0064), and may reasonably be interpreted as indicating a substantially greater WTP for clean air than that found in the original study.

Harrison, Rubinfeld (1978c, p. 87) use their estimation results to calculate a WTP value: "[W]hen NOX and the other values take on their mean values, the change in median housing values from a one pphm change in NOX is $1613". Table 3 shows similar values, obtained by taking the difference between mean predictions for the 96 and 489 observation data sets, and both data sets with a one pphm reduction in NOX. Note that the result for the OLS 489 observation data set is $1254, less than the original result for 506 observations. Mean predictions are used instead of predictions from mean values to accommodate the spatially lagged independent variables included in the SLX and SDEM models. The WTP value for the weighted SLX model estimated using the TASSIM zones data set is $4178, under the same assumptions as those in the original research, but derived from a model now without serious mis-specification problems most likely related to the choice of observational units.

4 Concluding remarks

It has been demonstrated that the difficulties experienced in extracting a reasonable estimate of the air pollution coefficient from a model of the Boston house value data set that is not strongly mis-specified may be resolved by changing the unit of observation to that of the air pollution variable. Using TASSIM model output zones as the appropriate unit of observation does reduce the observation count and consequently increases the uncertainty of the coefficient expressed by its standard error, but also removes most of the mis-specification issues. In concluding, it will be suggested that this is not a one-off issue with this particular data set, but fits into discussion of the choice of units of observation in Regional Science.

In *Complex spatial systems*, Wilson (2000) distinguishes three dimensions which interact in urban and regional analysis: system articulation, theory, and method (see also Wilson 2002, 2012). System articulation is, in turn, made up of three sub-dimensions: entitation, levels of resolution (sectoral, spatial, temporal), and spatial representation. He argues that all too little attention is paid in analysis to careful planning of the main dimensions, with system articulation typically treated in the least satisfactory way. His second and third chapters provide a succinct and enlightening review of why system articulation matters — pointing back to Paelinck, Nijkamp (1975).

Both the spatial level of resolution and the mode of spatial representation are involved in spatial scale (see also Dray et al. 2012). Scale is intimately connected to the pattern/process matching that is central to analysis, because certain causal effects may be present only at particular scales. If the spatial representation (driven by available data) misses this scale, mis-specification issues may emerge (see also McMillen 2003). By using TASSIM zone units of observation, chosen to match the spatial pattern of entities corresponding to the TASSIM model output zones, and because these largely also correspond to the town units for which many other independent variables were measured, we have in this case removed the very high levels of autocorrelation induced by "copying out" to multiple census tracts belonging to these entities.

This is related to but probably not a case of the modifiable areal unit problem (Gelfand 2010); the observations in the case of the census tract data set are misaligned because of duplication of data observed for a smaller number of spatial entities covering the same geographical area (see also Haining 2010). This induces very strong spatial autocorrelation between proximate neighbors but does not add any information; it is also likely that heteroskedasticity is also induced by duplication. There is also a relationship to the ecological fallacy, in that the discussion in Harrison, Rubinfeld (1978c) largely relates to household WTP, but no household-level data is available. Wakefield, Lyons (2010) give a survey of the ecological fallacy in connection with spatial aggregation; the point of concern is the extension of aggregated inference to individuals within the aggregates.

In conclusion, researchers should be aware of the impact that choices of units of observation will have on the results that they obtain. Had Harrison, Rubinfeld (1978c) chosen to use TASSIM model output zones as units of observation, they would not only have avoided the serious mis-specification that has been subsequently found in their OLS model, but they would also have been able to establish that WTP for clean air was about three times higher than they believed at the time. Understanding how spatial data is organized has not been paid sufficient attention in spatial econometrics, often, as in this case, leading to spurious spatial autocorrelation stemming more from the way the data has been handled than from underlying data generation processes.

Acknowledgements

The author would like to thank the editors and two anonymous reviewers for their contributions to this article. Achim Zeileis, Antony Unwin, Tony Smith and Jon Wakefield kindly commented on the first version of the paper. Further helpful comments have been received during presentations at the Bergen Economics of Energy and Environment Research Conference, the Spatial Econometrics Association X World Conference, the R User Conference, a seminar at Adam Mickiewicz University, and from participants at graduate courses in the Universities of Innsbruck and Tartu, as well as during the Workshop in Honor of Manfred M. Fischer's Retirement from his Professorship at WU Vienna. Remaining issues are entirely the responsibility of the author.

References

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.

Belsley DA, Kuh E, Welsch RE (1980) *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* John Wiley & Sons, New York. CrossRef.

Dray S, Couteron P, Fortin MJ, Legendre P, Peres-Neto PR, Bellier E, Bivand R, Blanchet FG, de Cáceres M, Dufour AB, Heegaard E, Jombart T, Munoz F, Oksanen J, Pélissier R, Thioulouse J, Wagner HH (2012) Community ecology in the age of multivariate multiscale spatial analysis. *Ecological Monographs* 82: 257–275. CrossRef.

Drukker DM, Egger P, Prucha IR (2013) On two-step estimation of a spatial autoregressive model with autoregressive disturbances and endogenous regressors. *Econometric Reviews* 32: 686–733. CrossRef.

Gelfand AE (2010) Misaligned spatial data: The change of support problem. In: Gelfand AE, Diggle P, Guttorp P, Fuentes M (eds), *Handbook of Spatial Statistics.* Chapman & Hall/CRC, Boca Raton, 517–539. CrossRef.

Gibbons S, Overman HG (2012) Mostly pointless spatial econometrics? *Journal of Regional Science* 52: 172–191. CrossRef.

Gilley OW, Pace RK (1996) On the Harrison and Rubinfeld data. *Journal of Environmental Economics and Management* 31: 403–405. CrossRef.

Gotway CA, Young LJ (2002) Combining incompatible spatial data. *Journal of the American Statistical Association* 97: 632–648. CrossRef.

Haining RP (2010) The nature of georeferenced data. In: Fischer M, Getis A (eds), *Handbook of Applied Spatial Analysis.* Springer, Heidelberg, 197–217. CrossRef.

Halleck Vega S, Elhorst JP (2015) The SLX model. *Journal of Regional Science* 55: 339–363. CrossRef.

Harrison D, Rubinfeld DL (1978a) The air pollution and property value debate: Some empirical evidence. *The Review of Economics and Statistics* 60: 635–638. CrossRef.

Harrison D, Rubinfeld DL (1978b) Distribution of benefits from improvements in urban air-quality. *Journal of Environmental Economics and Management* 5: 313–332. CrossRef.

Harrison D, Rubinfeld DL (1978c) Hedonic housing prices and the demand for clean air. *Journal of Environmental Economics and Management* 5: 81–102. CrossRef.

Ingram GK, Fauth GR (1974) *TASSIM: A Transportation and Air Shed SIMulation model, volume 1. case study of the Boston region.* Department of City and Regional Planning, Harvard University

Ingram GK, Fauth GR, Kroch EA (1974) *TASSIM: A Transportation and Air Shed SIMulation model, volume 2: program user's guide.* Department of City and Regional Planning, Harvard University

LeSage JP (2014) What regional scientists need to know about spatial econometrics. *Review of Regional Studies* 44: 13–32. CrossRef.

McMillen DP (2003) Spatial autocorrelation or model misspecification? *International Regional Science Review* 26: 208–217. CrossRef.

Ord J (1975) Estimation methods for models of spatial interaction. *Journal of the American Statistical Association* 70[349]: 120–126. CrossRef.

Pace RK, Gilley O (1997) Using the spatial configuration of the data to improve estimation. *Journal of the Real Estate Finance and Economics* 14: 333–340. CrossRef.

Pace RK, LeSage JP (2008) A spatial Hausman test. *Economics Letters* 101: 282–284. CrossRef.

Paelinck JHP, Nijkamp P (1975) *Operational theory and method in regional economics*. Saxon House, Farnborough

Piras G (2010) sphet: Spatial models with heteroskedastic innovations in R. *Journal of Statistical Software* 35: 1–21. CrossRef.

Solon G, Haider SJ, Wooldridge JM (2015) What are we weighting for? *Journal of Human Resources* 50: 301–316. CrossRef.

Wakefield JC, Lyons H (2010) Spatial aggregation and the ecological fallacy. In: Gelfand AE, Diggle P, Guttorp P, Fuentes M (eds), *Handbook of Spatial Statistics*. Chapman & Hall/CRC, Boca Raton, 541–558. CrossRef.

Waller LA, Gotway CA (2004) *Applied Spatial Statistics for Public Health Data*. John Wiley & Sons, Hoboken, NJ. CrossRef.

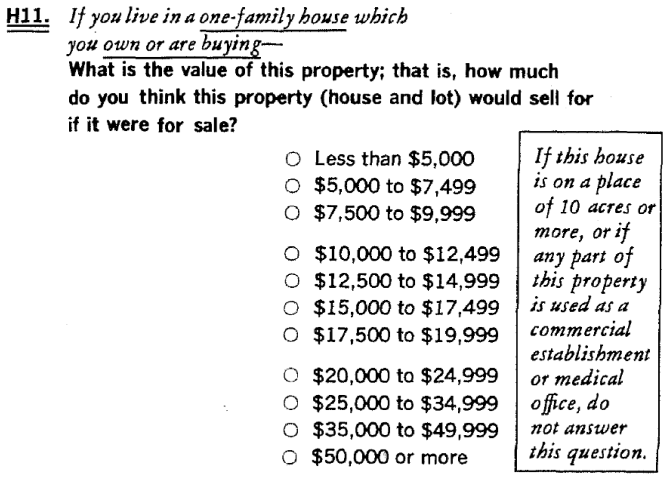
Wilson AG (2000) *Complex spatial systems: The Modelling Foundations of Urban and Regional Analysis*. Prentice Hall, Harlow. CrossRef.

Wilson AG (2002) Complex spatial systems: Challenges for modellers. *Mathematical and Computer Modelling* 36: 379–387. CrossRef.

Wilson AG (2012) *The science of cities and regions: lectures on mathematical model design*. Springer, Dordrecht. CrossRef.

Wong D (2010) The modifiable areal unit problem (MAUP). In: Fotheringham AS, Rogerson PA (eds), *The SAGE Handbook of Spatial Analysis*. SAGE Publications Ltd, London, 105–123. CrossRef.

A Appended figures and tables



H11. *If you live in a one-family house which you own or are buying—*

What is the value of this property; that is, how much do you think this property (house and lot) would sell for if it were for sale?

- ○ Less than $5,000
- ○ $5,000 to $7,499
- ○ $7,500 to $9,999
- ○ $10,000 to $12,499
- ○ $12,500 to $14,999
- ○ $15,000 to $17,499
- ○ $17,500 to $19,999
- ○ $20,000 to $24,999
- ○ $25,000 to $34,999
- ○ $35,000 to $49,999
- ○ $50,000 or more

If this house is on a place of 10 acres or more, or if any part of this property is used as a commercial establishment or medical office, do not answer this question.

Figure A.1: Scanned copy of 1970 Census question H11.

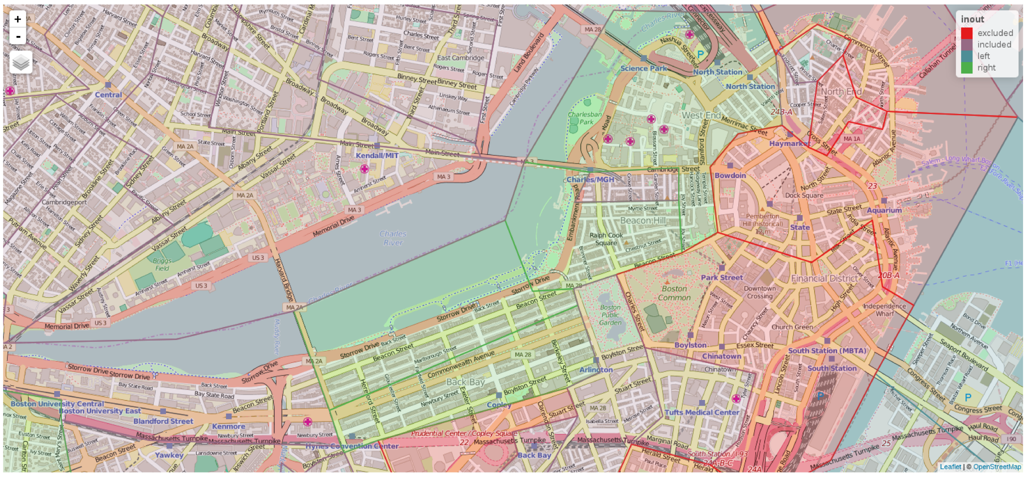


Figure A.2: A screen-dump of central Boston composed using the R interactive mapping function `mapview` shown on an OpenStreetMap basemap; red-shaded tracts are excluded from the 506 tract dataset, purple tracts are included, blue tracts are left-censored (<$5,000), and green tracts are right-censored (>$50,000).

Table A.1: Five-number summary of counts of block-based housing unit counts by block for reference, and of counts of blocks and aggregated block-based housing unit counts by tract.

	minimum	lower quartile	median	upper quartile	maximum
Block housing unit counts by tract	151.00	1174.50	1610.00	2165.00	5482.00
Block counts by tract	1.00	26.00	47.50	70.00	299.00
Block housing unit counts	-1.00	10.00	19.00	37.00	1564.00

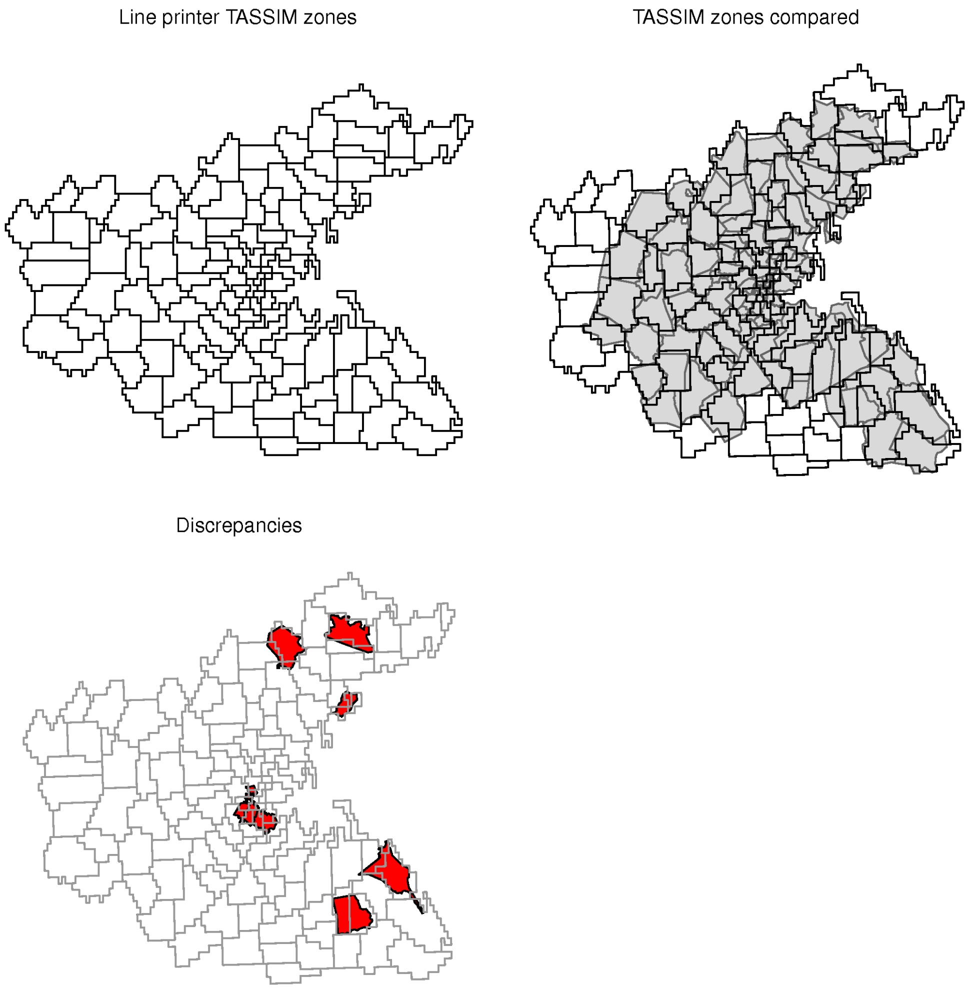


Figure A.3: TASSIM zone boundaries retrieved from line printer output compared with approximate TASSIM zones retrieved from 506 Census tract data set; upper left panel: boundaries of 122 line printer zones; upper right panel: affine-transformed line printer zones overlaid over 96 homogeneous and contiguous NOX zones from census tract dataset; lower left panel: discrepancies between the two boundary sets.

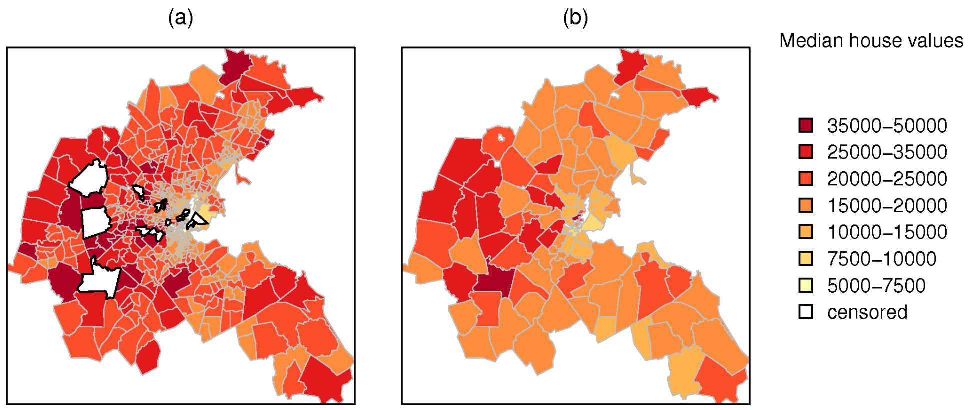


Figure A.4: Median house values (USD 1970): (a) 506 census tracts and (b) 96 approximate TASSIM zones; in panel (a), the censored tracts are shown with black boundaries.

Table A.2: Descriptives for variables included in the hedonic house value regression; 489 non-censored Boston census tracts

	Min.	Median	Mean	Max.
units	5.00	526.00	690.10	3031.00
log(median)	8.63	9.95	9.92	10.82
CRIM	0.01	0.25	3.45	88.98
ZN	0.00	0.00	11.13	100.00
INDUS	0.74	9.69	11.10	27.74
CHAS	0.00	0.00	0.06	1.00
I((NOX * 10)^2)	14.82	28.94	32.04	75.86
I(RM^2)	12.68	38.35	39.46	77.09
AGE	2.90	76.70	68.21	100.00
log(DIS)	0.13	1.19	1.20	2.50
log(RAD)	0.00	1.61	1.86	3.18
TAX	187.00	330.00	407.50	711.00
PTRATIO	12.60	19.10	18.52	22.00
I(BB/100)	0.00	0.01	0.06	0.96
log(I(LSTAT/100))	-3.92	-2.15	-2.21	-0.97

Table A.3: Descriptives for variables included in the hedonic house value regression; 96 approximate TASSIM zones

	Min.	Median	Mean	Max.
units	25.00	2926.00	3588.00	12410.00
log(median)	9.12	9.82	9.83	10.56
CRIM	0.01	0.08	1.96	18.13
ZN	0.00	0.00	25.89	100.00
INDUS	0.46	6.01	8.55	27.74
CHAS	-1.00	-1.00	-0.85	0.00
I((NOX * 10)^2)	14.82	21.76	26.68	75.86
I(RM^2)	25.93	39.60	41.82	62.77
AGE	8.97	51.76	56.02	100.00
log(DIS)	0.14	1.54	1.42	2.50
log(RAD)	0.00	1.61	1.62	3.18
TAX	187.00	307.00	376.20	711.00
PTRATIO	12.60	18.25	17.93	22.00
I(BB/100)	0.00	0.01	0.04	0.78
log(I(LSTAT/100))	-3.52	-2.46	-2.42	-1.43

Table A.4: Model AIC values for observations on 489 non-censored census tracts

	OLS	SLX	SEM	SDEM
Census tracts	-318.53	-405.93	-515.03	-563.15
Census tracts with weights	-512.77	-566.90	-684.34	-700.45

Table A.5: Model AIC values for observations on 96 approximate TASSIM zones

	OLS	SLX	SEM	SDEM
TASSIM zones	-95.68	-135.50	-97.20	-133.57
TASSIM zones with weights	-135.25	-151.84	-144.19	-150.04

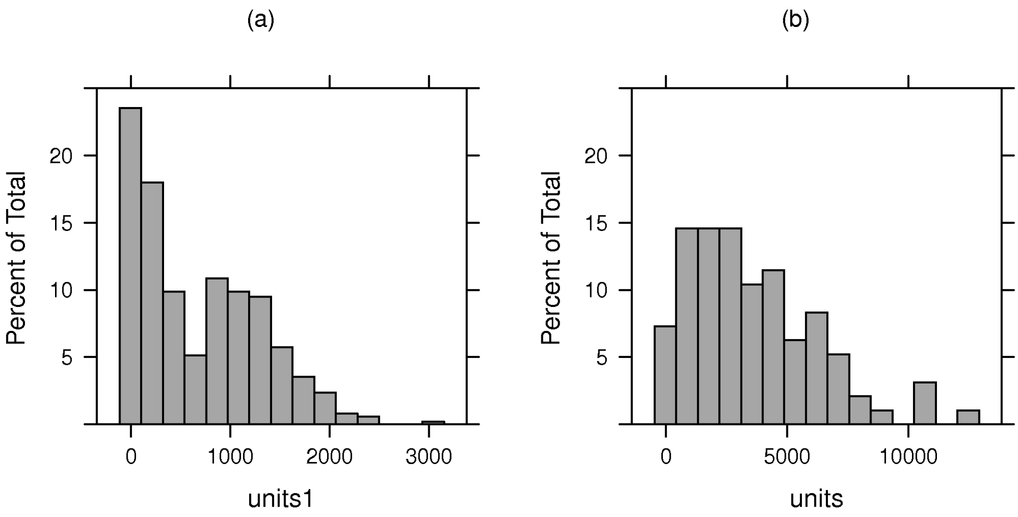


Figure A.5: Histograms of the distributions of counts of owner-occupied one-family housing units by observational unit; panel (a) observational units 506 Census tracts and panel (b) 96 approximate TASSIM zones.

Table A.6: Studentized Breuch-Pagan tests for heteroskedasticity for models fitted by least squares for the 489 census tract data set and the 96 TASSIM zone data set; columns 3 and 4 test against unit counts per aggregate entity as the source of heteroskedasticity.

	BP test	p-value	BP test with units	p-value
TASSIM zones OLS	51.429	1.6921e-06	4.942	0.02621
TASSIM zones SLX	29.901	0.27173	3.859	0.04949
Census tracts OLS	83.133	2.8297e-12	22.24	2.410e-06
Census tracts SLX	100.87	9.1962e-11	20.36	6.424e-06

Table A.7: Standard deviates of Moran's I test for spatial autocorrelation in least squares regression residuals for weighted and unweighted models and the 489 census tract data set and the 96 TASSIM zone data set.

	TASSIM zones	p-value	Census tracts	p-value
OLS	2.1767	0.01475	15.997	$< 2.2e-16$
Weighted OLS	4.3071	8.271e-06	13.901	$< 2.2e-16$
SLX	1.5739	0.05776	15.413	$< 2.2e-16$
Weighted SLX	1.9261	0.02704	13.21	$< 2.2e-16$

Table A.8: Robust Lagrange Multiplier tests for omitted spatial error and spatial lag terms in unweighted least squares regression models with and without the spatially lagged independent variables for the 489 census tract data set and the 96 TASSIM zone data set.

	Robust LM error	p-value	Robust LM lag	p-value
TASSIM zones OLS	0.10147	0.7501	5.0123	0.0251675
TASSIM zones SLX	0.093408	0.7599	0.05813	0.8094769
Census tracts OLS	234.4	$<2e-16$	11.362	0.0007496
Census tracts SLX	177.56	$<2e-16$	0.10978	0.7404000

Table A.9: Hausman tests for differences between least squares and spatial error an spatial Durbin model coefficients for weighted and unweighted models and the 489 census tract data set and the 96 TASSIM zone data set.

	TASSIM zones	p-value	Census tracts	p-value
SEM	31.113	0.005345	52.192	2.609e-06
Weighted SEM	19.027	0.163938	26.567	0.021907
SDEM	17.712	0.912070	53.272	0.001866
Weighted SDEM	19.199	0.862621	48.936	0.006024

Table A.10: Spatial coefficient ρ_{Err} values and Wald test p-values for spatial error and spatial Durbin error models for weighted and unweighted models and the 489 census tract data set and the 96 TASSIM zone data set.

	TASSIM zones	p-value	Census tracts	p-value
SEM	0.40436	0.001271	0.73273	$<$ 2.2e-16
Weighted SEM	0.46709	7.566e-05	0.69576	$<$ 2.2e-16
SDEM	0.056156	0.717637	0.65797	$<$ 2.2e-16
Weighted SDEM	0.09641	0.528010	0.62017	$<$ 2.2e-16

Table A.11: Model NOX coefficient values and impacts, displayed graphically in Figure 3 ($N-k$ standard errors in parentheses); the final column shows the WTP for a one pphm reduction in NOX also shown in Table 3.

Model type	Direct	Indirect	Total	WTP
TASSIM ZONE:NO WEIGHTS:OLS	-0.011799		-0.011799	2118
	(0.00269)		(0.00269)	
TASSIM ZONE:WEIGHTS:OLS	-0.0067082		-0.0067082	1157
	(0.002654)		(0.002654)	
TASSIM ZONE:NO WEIGHTS:SEM	-0.010464		-0.010464	1866
	(0.002909)		(0.002909)	
TASSIM ZONE:WEIGHTS:SEM	-0.0058313		-0.0058313	998
	(0.002906)		(0.002906)	
CENSUS TRACT:NO WEIGHTS:OLS	-0.0059344		-0.0059344	1254
	(0.001074)		(0.001074)	
CENSUS TRACT:WEIGHTS:OLS	-0.0030698		-0.0030698	640
	(0.001126)		(0.001126)	
CENSUS TRACT:NO WEIGHTS:SEM	-0.0024688		-0.0024688	520
	(0.001528)		(0.001528)	
CENSUS TRACT:WEIGHTS:SEM	0.00071832		0.00071832	-150
	(0.00155)		(0.00155)	
TASSIM ZONE:NO WEIGHTS:SLX	-0.013772	-0.019762	-0.033534	6693
	(0.002337)	(0.005609)	(0.005798)	
TASSIM ZONE:WEIGHTS:SLX	-0.010402	-0.011771	-0.022173	4178
	(0.002986)	(0.005925)	(0.005901)	
TASSIM ZONE:NO WEIGHTS:SDEM	-0.013755	-0.01947	-0.033225	6620
	(0.003242)	(0.007836)	(0.008189)	
TASSIM ZONE:WEIGHTS:SDEM	-0.010231	-0.011097	-0.021328	3999
	(0.004118)	(0.008241)	(0.008419)	
CENSUS TRACT:NO WEIGHTS:SLX	0.0027696	-0.01308	-0.010311	2223
	(0.002079)	(0.002461)	(0.001208)	
CENSUS TRACT:WEIGHTS:SLX	0.0038597	-0.0097492	-0.0058895	1240
	(0.002098)	(0.002468)	(0.001319)	
CENSUS TRACT:NO WEIGHTS:SDEM	0.0010074	-0.0099152	-0.0089077	1899
	(0.001682)	(0.002529)	(0.002239)	
CENSUS TRACT:WEIGHTS:SDEM	0.0030491	-0.0075263	-0.0044772	929
	(0.00173)	(0.002639)	(0.002469)	

REGION

The Journal of ERSA
Powered by WU

Volume 4, Number 1, 2017, 129–146
DOI: 10.18335/region.v4i1.146

journal homepage: region.ersa.org

Regional Public Stock Reductions in Spain: Estimations from a Multiregional Spatial Vector Autorregressive Model

Miguel A Márquez[1], Julian Ramajo[2], Geoffrey J D Hewings[3]

[1] University of Extremadura, Badajoz, Spain (email: mmarquez@unex.es)
[2] University of Extremadura, Badajoz, Spain (email: ramajo@unex.es)
[3] University of Illinois Urbana-Champaign, Urbana-Champaign, IL, USA (email: hewings@illinois.edu)

Received: 21 July 2016/Accepted: 12 May 2017

Abstract. The estimation of the impact of public investment on regional economic growth requires consideration of the spatio-temporal dynamics among the state variables of each region. According to recent data, public investment in Spain has fallen in recent years, and in some regions the reduction has been so severe that public investment has not been able to replenish depreciated capital. Recent austerity policies in Spain that feature temporary decreases in the accumulation of regional public capital should thus be evaluated in terms of their impact on the economy as a whole and on specific regions together with the spillover effects from one region to the rest of the regional system. This paper uses a multiregional integrated specification to model interdependencies across regions; our results indicate that while global decreases in public investment have a homogeneously negative effect on the output of all Spanish regions, these regions portray heterogeneous responses from localized public capital stock reductions over the simulation period considered.

Key words: Regional investment; Public capital stock; Spillovers; Multiregional; Spatial econometrics; Vector autoregressions; Spain

1 Introduction

The international financial and economic crisis has had a major impact on European national budgets with European governments now increasingly trying to rebalance budgets, in most cases, by considering substantial expenditure cuts to eliminate budget deficits. The regional impacts of these budget adjustments are unlikely to be spatially blind generating the possibility that they will exacerbate existing disparities in welfare. The impact of austerity measures on public capital stock needs to be evaluated carefully; restrictive fiscal measures will require the consideration of new strategies on the one hand, and on the other hand, the impacts that spending reductions will have on policies to stimulate economic growth. National governments need to consider how best to rationalize public investment, and new tools are necessary to guide and prioritize these investments. Most analysts agree that unless policies are refocused to take up the new challenges, GDP could potentially fall across many European countries. At the regional level, in the face of declining resources, nations will be forced to prioritize how and where they invest public funds across regional areas. In this context, it is necessary to develop and implement quantitative approaches to guide infrastructure development and this paper tries to make a contribution to the provision of strategic information. For example, from a welfare perspective, across the board cuts (i.e. public expenditures are reduced by a similar percentage in all regions) may have a different total impact from reductions that are targeted to specific programs or in specific regions – even though the total volume of expenditure cuts might be the same in both cases. The complexities of the spatial allocation and the manner in which spatial spillover effects move across the region system require formal analysis of the kind proposed in this paper[1].

It is well known that the empirical analysis of the aggregate effects of regional public capital provision requires spillover effects to be considered (see, for example, Boarnet 1998, Moreno, López-Bazo 2007). These spillover effects could play a key role in regional economic performance, since they can be inferred as externalities that could lead to increasing returns to scale. Indeed, public capital accumulated in one region can create additional growth potentials in other regions. In this paper, growth spillovers are understood to characterize a situation in which some of the growth in a region may be traced to the public investment in neighboring regions. Following Capello (2009), the transmission channels of the influences from changes in the capital stock in a single regional economy to the growth of neighboring regional economies (and vice versa) will be via trade linkages and market relationships. Further, it is expected that trade linkages will transmit either positive or negative effects.

In the literature, one can appeal to input-output based analyses such as the hypothetical extraction method of Dietzenbacher et al. (1993). More complete evaluations could be accomplished with a multiregional computable general equilibrium model, adopting a similar strategy but with the added benefit of being able to trace some of the economy-wide impacts of fiscal re-allocations and differential tax rates (see McGregor, Swales 2003). A third alternative, adopted in this paper, is to use a multiregional integrated spatial vector autoregressive model (see Ramajo et al. 2017) to capture both the short-run and longer-run impacts of significant changes in fiscal policy. None of these approaches by themselves would encompass all of the necessary perspectives to address the problems at hand, but together they offer a rich source of information for policy analysts. The present paper promotes the third approach as a complement to the other two (rather than as a substitute) with a particular focus on the evaluation of austerity measures.

According to a recent research (Serrano Martínez et al. 2017), public investment in Spain has fallen almost 60% since 2009 and in some regions the reduction has been so severe that it has not been able to replenish depreciated capital. As such, it would be very interesting to simulate the effects of a reduction in the public stock of infrastructures. This paper specifies a 'global' approach to assess how output in every Spanish region is affected by a reduction in the public capital stock of another specific Spanish region. Specifically, a multiregional integrated spatial vector autoregressive (MultiREG-SpVAR) model is applied that allows the investigation of spatio-temporal interdependencies across regions[2]. This model is used to assess the magnitude of both intra-regional (domestic) and inter-regional (spillover) effects of public infrastructure for the regions of Spain. Domestic effects are understood to be the effects derived from public capital installed in the region itself, while the spillover effects are those derived from public capital installed outside the region.

The rest of the paper is organized as follows. Section 2 briefly reviews some issues about both the domestic impact of public capital on regional economic growth and the

[1]Even here the full welfare effects will not be estimated; to provide a more complete accounting, embedding this type of analysis within a multiregional computable general equilibrium model would be necessary.

[2]The used model draws on the methodological contribution developed in Ramajo et al. (2017). The MultiREG-SpVAR model extends the traditional VAR models applied in the economics literature (see Kamps 2005) to assess the effects of public capital at aggregate national level. Thus, the proposed approach encompasses both the VAR methodology, that specifies the single-region temporal dynamics (see Márquez et al. 2011), the bi-regional VAR approach (see Márquez et al. 2010a), and the uni-regional Spatial VAR approach (Márquez et al. 2015), combining different region-specific VAR models in a global specification in which the state variables of each region are related to the state variables of the rest of the regions.

spillovers effects of this type of capital, providing the intuition behind the empirical findings. Section 3 exposes the econometric background behind the MultiREG-SpVAR specification. Section 4 describes the Spanish regional system, and presents the database. The empirical application in Section 5 estimates the domestic and spillover effects of public capital in the Spanish regional system, and a discussion of the main implications of the empirical results are provided. Finally, Section 6 provides some summary remarks.

2 Theoretical background

The effect of public-sector capital on economic growth has been a recurring issue in economics. Initiated the seminal article by Ratner (1983) more than three decades ago, this literature was further stimulated by Aschauer (1989), who hypothesized and tested the role of public infrastructure investment on productivity and produced the result that it had a substantial impact. Nevertheless, the posterior evidence was not homogeneous. To name only a few works, Lynde, Richmond (1992) and Morrison, Schwartz (1996) argued that public capital investments fosters economic growth, while the empirical work by Tatom (1991) and Holtz-Eakin (1994) among others found that public capital formation is not a robust determinant of private output; other authors like Seitz (1993) and Devarajan et al. (1996) found negative effects of public capital on economic growth.

There is a substantial literature in regional economics on the impact of public expenditure on regional economic growth (see, among others, Munnell 1990, Garcia-Mila et al. 1996, Sturm et al. 1998, Destefanis, Sena 2005). More recently, there is an emerging consensus in the literature about the positive impacts of public infrastructure on per-capita income and productivity in the private sector (see, for example, Henderson, Kumbhakar 2006, Cohen, Morrison Paul 2007, Heintz 2010), although the effects are not as large as Aschauer (1989)Aschauer (1989) suggested.

The review of the extensive literature by Romp, De Haan (2007) highlights different approaches used to estimate the effects of public capital on regional economic growth[3]: production functions, cost functions and VAR/VECM models. In a survey of the literature measuring the impact of public investment expenditure, De la Fuente (2010) focuses on studies estimating aggregate production functions or growth equations, concluding that the contribution of public infrastructure investment to productivity growth is significant. Within the cost function approach, the pioneering work of (Deno 1988) stimulated studies at the regional level (e.g. Seitz, Licht 1995, Morrison, Schwartz 1996); while the data requirements for this type of approach are greater than in the case of the production-function approach, the main advantage is that the cost-function approach is more flexible than the production-function approach.

Finally, among the main methods, the use of the VAR approach to test the significance of the dynamic effects of public capital on economic growth presents some advantages (Kamps 2005). Whereas other approaches assume a causal relationship running from the inputs to economic output, the VAR approach does not impose any causal links between the variables a priori[4]. The VAR approach, although atheoretical in nature, allows for the existence of indirect links between the variables under investigation. In addition, it provides the opportunity to test the number of long-run (cointegrating) relationships. Moreover, if the number of long-run (cointegrating) relationships are tested and estimated consistently, the vector error correction (VEC) models would produce consistent estimates of the impulse response functions. With respect to the empirical literature where the VAR methodology has been used to simulate the effects of unexpected changes in the public capital on macroeconomic variables, it is possible to find different contributions at both country and regional levels. Recently, for example, Hunt (2012) worked with a set of OECD countries, Deliktas et al. (2009) analyzed seven regions of Turkey and Roca-Sagalés, Sala (2006) investigated the effects of public capital using the VAR approach for seventeen Spanish regions. Much of the previous empirical work appears to posit the existence of cross-sectional independence, which is unlikely to be the case.

[3]Considering different perspectives, these authors systematized previous research about the role of public capital in economic growth.

[4]Therefore, we will provide report results estimates according to which endogeneity of regressors (particularly public capital) is not a problem.

Although the approach used in the present paper shares the same starting point with previous empirical work using the VAR approach, the present approach relates more closely in spirit to recent work in regional economic growth through its preoccupation with the problem of spatio-temporal dynamics (e.g. Márquez et al. 2010b). Clearly, the processes of regional economic growth intrinsically embrace space and time interactions. Even though the effects of public capital formation considering spatio-temporal interactions at the regional level have received increasing attention over recent years (for example, see Pereira, Roca-Sagalés 2003), past research on the impact of public capital on regional growth has mainly ignored two relevant challenges. First, it has not considered the estimation of the spillover effects of a reference region's public capital on the economic growth of any other region of the country. Secondly, there has been no attempt to integrate, in some fashion, the state variables of each region with the state variables of the rest of the regions to provide a global perspective that allows for the investigation of interdependencies across regions. Consequently, the resulting parameter estimates and statistical inference of the empirical work produced to date could be open to criticism. Tackling these two problems simultaneously will allow the quantification of the contribution to overall regional economic growth of the provision of public capital in a determined region, making available an important tool to assist in formulating economic policies.

To deal with these open questions, the approach used in the present paper is the confluence of the recent empirical literature on the effects of public capital using VAR models (consistent with the argument that the analysis of these effects requires the consideration of dynamic feedbacks among the different variables) with the new developments in macroeconometric modeling following the Global VAR approach proposed by Pesaran et al. (2004) and Dees et al. (2007). The multiregional integrated approach (Ramajo et al. 2017) will contemplate both the spatial and temporal growth output effects derived from public infrastructure at the regional level.

At the aggregate level, it is not easy to disentangle the effects of public expenditure on regional growth; it is especially difficult to identify the channels through which they operate. The economics literature (see among others Gramlich 1994, Romp, De Haan 2007) highlights different channels through which public investments influence economic growth. The first channel makes reference to the direct increased flow of output due to the accumulation or the stock of public capital over time. Other channels are connected to the effects derived from the interactions between public and private capital; namely, the direct effects on productivity of private production inputs, the crowding-out effect on private spending through financial mechanisms, and the complementary effect from private investment. Following Baxter, King (1993), the aforementioned channels could generate both short-run and long-run effects. These traditional effects operate at the same time as other effects emphasized by Agénor, Moreno-Dodson (2006): indirect effects on the productivity of workers, effects related to adjustment costs in both private capital formation and its mobility, effects connected to the creation of production facilities that encourage economic activities, effects derived from enhancing the durability of private capital, and effects resulting from the positive outcomes on growth coming from improvements in health and education. Hence, there are a variety of ways in which public capital could affect economic growth.

It is important to highlight that the presence of strong spatio-temporal dynamic interaction effects can alter the standard predictions about the macroeconomic implications of regional public investment. The shifts in public capital could lead to potential multiplier effects that induce dynamic responses on regional growth. In turn, these dynamic responses within a single regional economy can influence the growth of neighboring regional economies (Cheshire, Carbonaro 1996). Although different works provide arguments that consider spatial spillovers from public infrastructures (Holtz-Eakin, Schwartz 1995, Álvarez et al. 2006, Yu et al. 2013)(see among others, Holtz-Eakin and Schwartz, 1995; Alvarez et al., 2006; Yu et al., 2013), regional trade linkages occupy a relevant place among the different transmission channels of the effects of regional public capital expenditure on the growth of the regions integrating the regional economic system (Capello 2009). Following the methodological approaches presented in Márquez et al. (2010a), a distinction can be made between growth spillovers sent from a region to its trade-related neighboring regions (push-out effects) and growth spillovers received for a region from its trade-related neighboring regions (push-in effects). From a theoretical perspective, push-in and push-out effects could exert positive or negative influences on the growth of regional economies (Young 1991, Chung, Hewings 2015). Nevertheless, and according to the recent literature (see Melitz, Ottaviano 2008), one could expect that regional growth spillovers transmitted through trade linkages would have a positive effect on the growth of other regional economies in the short run; but the impact in the long-run may be ambiguous. Effectively, these growth spillovers could stimulate regional growth through, among other ways, increases in productivity derived from competitive pressures from trade, from learning-through-trade, and from enhanced access (lower cost, more reliable service etc.) to inputs, outputs and intermediates. Alternatively, trade can displace regional production and regional factor productions, leading to a negative effect on growth. Taken all together, it is clear that the empirical outcomes derived from the effects of regional public capital on the growth of trade-related regional economies can be mixed. All of these growth spillovers must be addressed in advocating and developing an alternative to current approaches to measuring and understanding what are the regional effects of public investment.

3 The Multiregional Spatial Vector Autoregressive (MultiREG-SpVAR) model[5]

To capture both the domestic and the spillover effects of public capital among the Spanish regions, a multiregional integrated spatial vector autoregressive model has been built following the Global VAR macroeconometric modeling approach proposed by Pesaran et al. (2004) and Dees et al. (2007).

In similar fashion to Global VAR models, our MultiREG-SpVAR specification provides a multi-location framework that allows the investigation of interdependencies across regions and is composed of individual models formulated by corresponding spatial VAR specifications (namely VAR models augmented by spatially-lagged variables) that are combined in a consistent manner. Each region is linked with the others in the regional system under study by including spatial lags in the econometric specification. In this way, all regions are potentially affected by developments in the other regions of the system. In this paper, the spatial lags have been constructed using spatial weights that are based on region-region trade flows as explained in Section 4.

Consider N regions, indexed by $i = 1, 2, \ldots, N$, then the Spatial VAR (SpVAR) model for each region i at time t ($t = 1, 2, \ldots, T$) is formulated as follows:

$$\begin{aligned} \boldsymbol{x}_{it} &= \boldsymbol{\Lambda}_{0i} + \boldsymbol{\Lambda}_{1i}\boldsymbol{t} + \boldsymbol{\Gamma}_{1i}\boldsymbol{x}_{i,t-1} + \ldots + \boldsymbol{\Gamma}_{p_i i}\boldsymbol{x}_{i,t-p_i} + \\ &\quad \boldsymbol{\Phi}_{0i}\boldsymbol{x}^*_{it} + \boldsymbol{\Phi}_{0i}\boldsymbol{x}^*_{it} + \boldsymbol{\Phi}_{1i}\boldsymbol{x}^*_{it-1} + \ldots + \boldsymbol{\Phi}_{q_i i}\boldsymbol{x}^*_{it-q_i} + u_{it} \end{aligned} \quad (1)$$

where $\boldsymbol{x}_{it}$ is the vector of internal state variables, $\boldsymbol{x}^*_{it}$ is a vector of external spatially-lagged variables that summarizes the state of the economy in the other regions (their components, $x^*_{g,it} = \sum_{j=1}^{N} w_{ij} x_{g,jt}$, are a weighted average of x_g in all regions except the i^{th}, since, by convention, $x_{ii} = 0$), $\boldsymbol{t}$ is a vector of deterministic time trends, $\boldsymbol{\Lambda}_{jt}$ $(j = 0, 1)$, $\boldsymbol{\Gamma}_{ji}$ $(j = 1, 2, \ldots, p)$ and $\boldsymbol{\Phi}_{ji}$ $(j = 0, 1, 2, \ldots, q)$ are matrices of coefficients to be estimated, and u_{it} is a vector of disturbances assumed to be serially uncorrelated with a zero mean and a non-singular covariance matrix, $\boldsymbol{\Sigma}_{ii}$.

The first step in building the MultiREG-SpVAR model, each regional SpVAR model is estimated individually and then the individual models are stacked to yield a 'global' compact specification in terms of a vector containing all the endogenous variables of the regional system, $\boldsymbol{x}_t = (\boldsymbol{x}^{'}_{1t}, \boldsymbol{x}^{'}_{2t}, \ldots, \boldsymbol{x}^{'}_{Nt})^{'}$:

$$\boldsymbol{x}_t = \boldsymbol{\Psi}_0 + \boldsymbol{\Psi}_1\boldsymbol{t} + \boldsymbol{F}_1\boldsymbol{x}_{t-1} + \ldots + \boldsymbol{F}_r\boldsymbol{x}_{t-r} + \boldsymbol{e}_t \quad (2)$$

[5]This section draws on Ramajo et al. (2017), where a more detailed technical explanation of the underlying econometric modeling approach is made.

for appropriate matrices $\boldsymbol{\Psi}_j$ $(j = 0, 1)$ and $\boldsymbol{F}_k$ $(k = 0, 1, \ldots, r)$, with $r = \max(p_i, q_i)$, where there are no restrictions on the covariance matrix $\Sigma_e = Cov(\boldsymbol{e}_t)$. This expression is the basis for the analysis of the dynamic properties of the multiregional mode.

Although the model allows consideration of the regional interdependencies and heterogeneity in the underlying dynamic process, analyzing within the regional system the response of every region to a temporary shock in a state variable of one specific region (for more details, see Ramajo et al. 2017), the model does not allow the quantification of the relevance of different channels for transmission of spillovers.

4 The Spanish regional system and the database

The database used consists of yearly time series over the period 1964-2003 for the Spanish Autonomous Communities[6]. The analysis of this specific period is the interest of the paper. The paper explores the space-time dimensions of this evolutionary period, and the Multireg-SpVAR model is appropriate to accomplish this. The Autonomous Communities of Spain have achieved the status of self-governed territories, sharing governance with the Spanish central government within their respective territories. Given the transfer of important economic responsibilities from the central government to the regional executives, the NUTS2 level of disaggregation in Spain is the most interesting level from a political economy perspective. Currently, almost one third (about 31%) of the regional investments are carried out by the autonomous communities, almost a quarter (about 24%) by the central government, with local corporations and other public entities accounting for the remaining 45%. Thus, both central government and regional authorities are able to decide how public expenditures are allocated.

For each region, the macroeconomic variables that compose the vector $\boldsymbol{x}_{it}$ are the following: gross value added, GVA, measured at basic prices in thousands of year 2000 constant euros; total employment (E), in thousands of employed persons; and private (KPR) and public (KPU) net capital stocks[7], in thousands of year 2000 constant euros. The regional series for GVA and E have been drawn from the *BD.MORES* database (Bustos et al. 2008) and the time series for KPR and KPU have been taken from the *Fundación BBVA-Ivie* database (Mas et al. 2009)[8]. Some summary statistics of these variables for the seventeen Spanish regions are shown in Table 1 (for reasons of comparability, the figures in this table are percentage shares; however, all the calculations in the paper have been made with the original data in logarithmic terms).

With respect to the vector of non-domestic variables $\boldsymbol{x}^*_{it}$, in our application, these have been built using trade-based weights w_{ij} in order to capture the economic interaction of region j with the i^{th} region's economy and not only the geographic interaction. These weights were computed using data on interregional trade in Spain drawn from the *C-Intereg* database (Llano et al. 2008, 2010). Initially, mean trade shares ($\bar{s}_{ij}$) were computed as the proportion of region j in the total trade (exports plus imports) of region i over the period 2004-2007 (measured in millions of euros)[9]. Thereafter, a binary trade-based spatial weight matrix was built defining non-normalized weights as and finally a row-standardized weights matrix was defined as

[6] This corresponds with a NUTS2 level of disaggregation in Eurostat nomenclature of statistical territorial units.

[7] The regional public capital stock comprises productive public capital owned by the local, regional, and national administrations, including transport infrastructure (roads, ports, airports, and railways), water and sewage facilities, and urban structures. As indicated by a referee, the analysis would benefit from a distinction between transport public infrastructures and the rest of productive infrastructures.

[8] The estimates of capital stock for the Spanish economy meet the methodological recommendations of the OCDE and converge with the procedure followed by the Bureau of Labor Statistics (BLS) and the EU KLEMS project. All the details about methodological aspects and the construction and revision of the database along time can be found in Mas et al. (2011).

[9] The out-of-sample period 2004-2007 was used to avoid endogeneity problems with the in-sample values of the state variables. The interregional trade flows data for Spain show a robust empirical regularity for the Spanish trade pattern (see Llano et al. 2010), demonstrating the existence for each Spanish region of stable relationships in terms of clients and suppliers. In essence, regional trade patterns do not change much over time.

Table 1: Summary statistics of the database (measured by the percentage of share)

	GVA	**E**	**KPR**	**KPU**
AND	13.8	15	13.6	14.9
ARA	3.1	3.2	3.5	4.4
AST	2.2	2.2	2.5	3.1
BAL	2.4	2.4	2.7	1.7
CAN	4.1	4.1	4.5	4.3
CANT	1.3	1.3	1.3	1.5
CAT	18.8	17.8	18.4	14
CLM	3.4	4	3.8	4.9
CYL	5.5	5.7	6.2	8.1
EXT	1.7	2	2	2.5
GAL	5.1	5.7	5.2	6.4
MAD	17.7	15.8	15.6	13.7
MUR	2.5	2.8	2.3	2.3
NAV	1.7	1.7	1.8	1.9
PV	6.2	5.5	6.4	6
RIO	0.8	0.8	0.8	0.9
VAL	9.7	10.3	9.4	9.3
SPAIN	**100**	**100**	**100**	**100**

Source: Own elaboration from BD.MORES and Fundación BBVA-Ivie databases.
Regional Abbreviations: Andalusia (AND), Aragón (ARA), Asturias (AST), Balearic Islands (BAL), Canary Islands (CAN), Cantabria (CANT), Catalonia (CAT), Castile-La Mancha (CLM), Castile and León (CYL), Extremadura (EXT), Galicia (GAL), Madrid (MAD), Murcia (MUR), Navarre (NAV), Basque Country (PV), La Rioja (RIO) and the Valencian Community (VAL).

$$w^*_{ij} = \begin{matrix} 0 & \text{if } i = j \\ 1 & \text{if } \bar{s}_{ij} \geq 0.01 \\ 0 & \text{if } \bar{s}_{ij} < 0.01 \end{matrix} \tag{3}$$

and finally a w row-standardized weights matrix was defined as $w_{ij} = w^*_{ij}/\sum_j w^*_{ij}$.

Thus, only trade-neighbors of region i those regions j that have a mean trade share above 10% of the total trade were used. This criterion is based on the idea that only 'relevant trader regions' have non-negligible spatio-temporal effects on their neighbors; the remaining regions were assumed to be less important and thus have negligible individual impacts. The 10% critical cut-off point yields a set of 3 to 5 (not necessarily contiguous) neighbors. To check the robustness of our results, two other values (5% and 15%) were adopted, and the results were qualitatively similar to those presented in this paper, and thus were omitted for the sake of brevity.

5 Empirical application: effects of public capital in the Spanish regional system

The issue of the effects of public capital formation on economic development is currently at the center of the policy debate in many European countries. The policies available to states are limited because, given existing resources, it is not possible to maintain the current levels of public investment. Due to economic and budget pressures, nations have to prioritize where and how they allocate public expenditure.

This section focuses on the estimation of both domestic and spillover effects of public capital innovations on regional outputs in Spain. The effects derived from a hypothetical reduction of public capital in the Spanish regional system are identified in order to determine the key regions that generate, sustain and improve the productive capacity of the regional economic system. The detection of these 'key regions' would contribute to enhancing the set of strategic information necessary to ensure the most efficient allocation of regional public investment in the context of economically interrelated regions and in prioritizing the provision of public resources.

Next, the estimation results of the MultiREG-SpVAR for the Spanish regional system are presented. Due to the very large number of regressions, as well as preliminary or intermediate results involved in the estimation, only the main estimation and specification test results are presented (see Ramajo et al., 2015, for details). In summary, given that all the conditions for the Global VAR approach are accomplished, the estimation of the MultiREG-SpVAR is justified, and the dynamic properties of the model can be investigated[10].

Thus, the purpose is to simulate the response of the system to either a unitary (one standard error) shock in one internal variable in one specific region, or a unitary global shock (a GVA-weighted average of variable/regional specific unit shocks across all the regions in the system) in the whole regional system. Given the existence of trade links between regions that are incorporated in the model, other regions in the system will be (more or less) affected from the disturbance, providing relevant information about the degree of interregional spillovers in the Spanish regional system.

In this application, both the domestic and spillover effects of one-time innovations in the public capital installed in one region are estimated, as are the effects of a global shock to public capital in the country on the output growth of the regions of Spain. Given the actual economic situation of the Spanish economy, characterized by a sizeable reduction in the government's current and infrastructure expenditure, negative shocks to public capital are simulated as either area-wide or individual shocks scenarios: a) a negative global shock to total public capital in Spain; and b) a negative unit shock to public capital in each region.

Tables 2 and 3 show detailed estimated effects in the first and the fifth years. In these tables, the last row shows the effects of a global shock, not originating in a particular region but common to the Spanish economy as a whole (it is defined as a GVA-weighted average of shocks to the public capital in all 17 regions in Spain); the remaining rows display the region-specific shocks, where the shaded cells document the own-response of output to a domestic public capital shock, and the non-shaded elements record the spillover effects in the different regions of Spain. As can be seen in Tables 2 and 3, the responses of output to a negative unit shock in public capital have more heterogeneous responses (positive and negative effects) initially than after five years (where negative responses clearly prevail).

Starting from the first shock scenario (a negative global shock to total public capital in Spain), the effects on output of a common global negative shock to the public capital in the Spanish economy as a whole are contained in the last row of Tables 2 and 3. It should be noted that the term global is used in a way that departs from the empirical definition applied in other empirical research. In this paper, the term global shock makes reference to an overall shock for the Spanish regional system, including all of the dynamic feedbacks among the trade-related regions.

While a common global negative shock to the Spanish economy as a whole has initially mixed effects on the regions (positive and negative, see Table 2), after 5 years the shock has similar (negative) impacts (see Table 3). The strongest negative effect from this initial (negative) global shock can be found in Castile La Mancha (-2.49%), closely followed by the Basque Country (-2.37%). The weakest negative effects are found in La Rioja (-0.64%), Extremadura (-0.65%) and Cantabria (-0.93%).

Generally, from Table 3, the GVA-weighted average shock shows responses of GVA that are similar to the ranking in the share of national GVA (the greater the share, the more negative the impact; see Figure 1). Some exceptions are the regions that have the largest negative effects, Castile-La Mancha (ninth position within the Spanish ranking considering its share in national GVA) and the Basque Country (fifth position in GVA shares). For Castile-La Mancha, the reason could be the lack of agglomeration derived

[10]The dynamic analysis presented in this section follows the Generalized Impulse Response (GIR) approach proposed by Koop et al. (1996), and developed further in Pesaran, Shin (1998) for VAR models. This approach generalizes the traditional Orthogonalized Impulse Response method of Sims (1980), being invariant to the ordering of the variables in the SpVAR models, which makes the GIR functions a very useful tool to analyze the propagation of shocks across regions.

Table 2: Estimated contemporaneous responses of output to a negative unit shock in public capital

	Response (%) in:																
Shock in:	**AND**	**ARA**	**AST**	**BAL**	**CAN**	**CANT**	**CAT**	**CLM**	**CYL**	**EXT**	**GAL**	**MAD**	**MUR**	**NAV**	**PV**	**RIO**	**VAL**
AND	-0.21	-0.05	-0.04	-0.21	0.31	0.06	-0.1	0	-0.09	0.15	0.15	0.07	-0.06	0.05	-0.05	0.04	0.06
ARA	-0.01	0.07	-0.15	0.05	0.62	0.09	0.05	-0.01	0.04	0.27	0.25	0.02	0.17	0.05	-0.15	0.09	0.16
AST	0.01	-0.05	-0.05	-0.02	0.11	-0.16	-0.09	0.24	0.03	-0.04	0.1	-0.18	-0.04	0.17	-0.22	0.01	-0.09
BAL	-0.02	0.14	0.28	0.07	-0.16	0.1	0.02	0.08	0.05	0.16	-0.02	0.09	-0.08	-0.1	0.28	0.36	-0.05
CAN	0.02	-0.03	-0.14	-0.28	-0.02	-0.11	-0.06	0.11	0.05	0.29	-0.09	-0.1	0.09	-0.2	-0.17	0.04	-0.09
CANT	-0.23	-0.22	0.05	-0.1	0.15	-0.14	-0.31	-0.05	0.02	0.11	0.06	-0.18	-0.17	-0.13	-0.08	0	-0.21
CAT	0.03	0.01	-0.12	0	-0.09	-0.1	0.09	-0.06	0	0.04	-0.02	-0.14	0.16	0.04	-0.19	-0.23	-0.02
CLM	0.14	0.13	0.02	0.26	0.06	-0.16	-0.03	-0.09	0.24	0.18	-0.09	-0.02	0.33	-0.29	0.08	-0.17	0.02
CYL	0.13	0.16	0.01	0.11	0.29	-0.02	0.04	0.19	0.18	0.18	0.16	0	0.11	0.03	-0.18	0.15	0.12
EXT	0.04	0	0.1	0.36	0.28	0.14	-0.02	-0.03	0.06	-0.04	0.11	-0.08	-0.08	-0.01	0.06	0.08	-0.03
GAL	0.01	0.06	-0.15	-0.17	0.24	0.07	0.1	-0.1	0.03	0.05	-0.02	0.12	0.1	-0.12	-0.04	-0.1	-0.18
MAD	-0.08	0.15	-0.14	-0.06	-0.14	0.05	0.08	-0.21	-0.1	-0.05	-0.11	0.09	0.15	0.18	-0.13	0.09	0.16
MUR	-0.23	0.02	-0.02	-0.13	0.21	0.06	-0.26	-0.21	-0.01	0.46	0.06	-0.39	-0.24	-0.43	-0.38	0.1	-0.13
NAV	-0.08	-0.01	-0.19	-0.06	0.51	0.1	-0.09	-0.12	-0.02	0.21	0.2	-0.19	-0.2	-0.09	-0.2	0.14	0.06
PV	-0.13	0.13	-0.06	-0.15	-0.15	0.19	0.09	-0.12	-0.05	-0.05	-0.07	-0.1	-0.09	-0.15	-0.3	-0.17	-0.13
RIO	-0.11	-0.1	0.2	0.18	-0.01	0	-0.02	-0.17	-0.13	0.13	-0.02	0.16	-0.02	-0.12	0.16	0.01	-0.04
VAL	-0.08	-0.12	-0.16	-0.22	0.01	-0.1	0.11	-0.1	-0.2	-0.16	-0.16	0.19	0.03	0.07	0.05	-0.21	-0.11
GLO	**-0.23**	**0.14**	**-0.33**	**-0.28**	**0.17**	**0.03**	**0.08**	**-0.32**	**-0.16**	**0.21**	**-0.04**	**0**	**0.2**	**0.04**	**-0.41**	**-0.02**	**0.07**

Notes: Numbers reported are median estimates of initial responses (impact) of GVA to one standard error negative shock to public capital in the region row or one standard error negative global shock to public capital. The bootstrapping method provides the impact on median bootstrap estimates (see the supplement A of Dees et al. 2007).

Table 3: Estimated responses of output to a negative unit shock in public capital after five years

	Response (%) in:																
Shock in:	**AND**	**ARA**	**AST**	**BAL**	**CAN**	**CANT**	**CAT**	**CLM**	**CYL**	**EXT**	**GAL**	**MAD**	**MUR**	**NAV**	**PV**	**RIO**	**VAL**
AND	-0.53	-0.08	-0.21	-0.81	0.25	0.33	-0.43	-0.64	-0.41	-0.19	-0.08	-0.07	-0.17	0.02	-0.26	0.3	0.17
ARA	-0.87	-0.38	-0.63	-0.69	0	-0.15	-0.94	-1	-0.48	-0.31	-0.28	-0.77	-0.7	-0.62	-0.89	-0.05	-0.44
AST	-0.45	-0.42	-0.44	-0.34	-0.55	-0.63	-0.7	-0.29	-0.19	-0.29	-0.34	-0.79	-0.6	-0.48	-0.83	-0.19	-0.66
BAL	-0.42	-0.19	-0.02	-0.03	-0.51	-0.29	-0.5	-0.37	-0.15	-0.11	-0.24	-0.43	-0.53	-0.54	-0.08	0.15	-0.53
CAN	0.11	0.12	-0.12	-0.24	0.13	0.11	0.11	0.19	0.04	0.36	-0.05	0.15	0.23	0.17	0.07	0.14	0.11
CANT	-1.27	-1	-0.79	-0.79	-0.81	-0.99	-1.61	-1.4	-0.61	-0.65	-0.71	-1.34	-1.24	-1.16	-1.36	-0.41	-1.12
CAT	-1.08	-1.08	-1.16	-0.73	-1.53	-1.29	-1.37	-1.26	-0.56	-0.37	-1.01	-1.67	-1.2	-1.45	-1.85	-1.01	-1.51
CLM	0.04	-0.02	-0.06	0.26	-0.11	-0.37	-0.11	-0.37	0.15	-0.03	-0.14	-0.21	0.06	-0.38	-0.03	-0.22	-0.26
CYL	-0.32	0.11	-0.03	-0.37	0.18	0.19	-0.4	-0.51	-0.08	-0.21	0.09	-0.22	-0.28	-0.13	-0.27	0.34	0.11
EXT	0.47	0.31	0.46	0.78	0.74	0.31	0.52	0.5	0.27	-0.05	0.42	0.38	0.28	0.45	0.64	0.14	0.22
GAL	-0.23	-0.06	-0.31	-0.37	0.05	0.09	-0.13	-0.28	-0.12	-0.1	-0.06	-0.07	-0.09	-0.13	-0.2	-0.14	-0.35
MAD	-0.64	-0.39	-0.61	-0.47	-0.93	-0.39	-0.64	-0.91	-0.37	-0.16	-0.58	-0.7	-0.62	-0.65	-0.84	-0.13	-0.6
MUR	-0.86	-0.52	-0.65	-0.52	-0.46	-0.41	-1.03	-1	-0.48	-0.31	-0.51	-1.16	-0.88	-0.92	-1.12	-0.17	-0.76
NAV	-0.68	-0.33	-0.55	-0.59	0.16	-0.06	-0.78	-0.89	-0.41	-0.43	-0.15	-0.74	-0.72	-0.44	-0.78	0.11	-0.24
PV	0.19	0.42	0.26	-0.04	0.01	0.62	0.44	0.3	0.16	0.17	0.18	0.25	0.29	0.28	0.48	0	0.13
RIO	0.24	0.28	0.6	0.41	0.56	0.38	0.52	0.27	0.05	0.13	0.37	0.73	0.5	0.48	0.78	0.11	0.51
VAL	-0.52	-0.2	-0.36	-0.73	-0.24	0.02	-0.31	-0.59	-0.32	-0.28	-0.23	-0.06	-0.24	-0.19	-0.28	-0.19	-0.17
GLO	**-2.09**	**-1.17**	**-1.67**	**-1.83**	**-1.78**	**-0.93**	**-2.2**	**-2.49**	**-1.12**	**-0.65**	**-1.39**	**-2.06**	**-1.79**	**-1.76**	**-2.37**	**-0.64**	**-1.72**

Notes: Numbers reported are median estimates of responses of GVA to one standard error negative shock to public capital in the region row or one standard error negative global shock to public capital. The bootstrapping method provides the impact on median bootstrap estimates (see the supplement A of Dees et al. 2007).

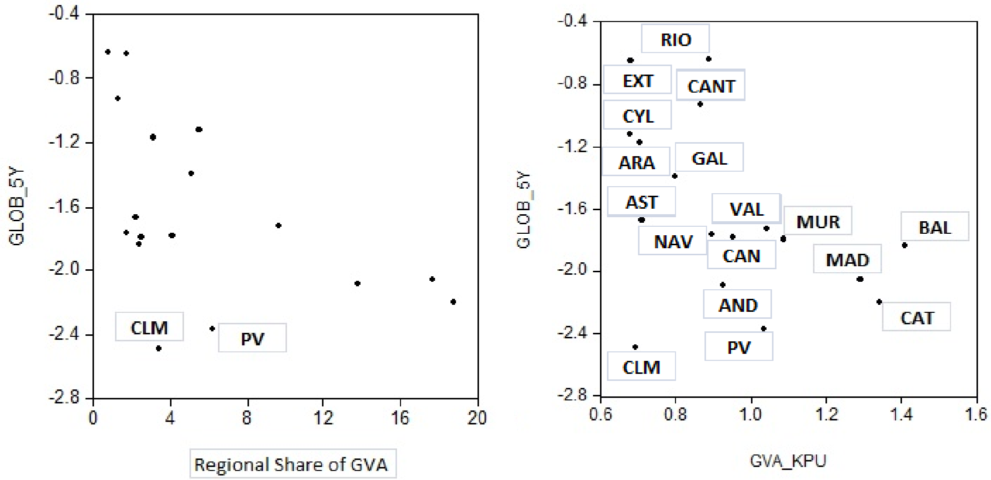


Figure 1: Effects after 5 years of a global negative shock (GLOB_5Y) vs. regional shares of national GVA and the regional ratios GVA/public capital (GVA_KPU)

from the low density that characterizes this region while in the case of the Basque Country, the explanation may lie in its economic and institutional differences from the rest of Spain.

On the other hand, Figure 1 also shows a negative correlation between the economic impact of a negative shock to public capital and the ratio of GVA to public capital. It seems that a negative global shock to public capital will induce higher negative response on GVA when the GVA/Public Capital ratio increases. This negative correlation seems to have little effect on the region Castile-La Mancha, since this region has an existing low ratio and it presents the lowest effect.

Overall, the estimation results suggest that a negative global shock will have a negative effect on the output of every region in Spain. However, global shocks can hide the detection of locations where the spillovers originate. Further, operating only with a global shock does not allow isolation of the effects from a specific region of interest. Consequently, the second shock scenario (a negative unit shock to public capital in each region) is needed in order to identify the regions where further reduction in public capital would have a greater negative impact, a lower negative impact, or even positive effects. Estimations of the domestic and the spillover output effects (from a region on the rest of the system) of one standard error negative shock to public capital in every Spanish region are provided in Tables 2 and 3. These results are new in the literature, since previous studies have only been able to reveal a part of the complex spatio-temporal feedbacks across regions; none of these studies has identified the spatio-temporal interdependencies between the state variables of all the regions.

From the estimates contained in Tables 2 and 3, the domestic and the spillover effects display positive and negative signs. Although negative effects prevail, the existence of positive effects opens a new dimension for the analysis of public investment. As seen in Table 3, for three of the seventeen regions, the domestic effects after five years are positive. Reduced funding for public capital accumulation in the Basque Country, Canary Islands and La Rioja will affect positively their regional economic growth. The underlying explanation might be in the fact that, while public capital formation will raise output directly in these regions (public and private capital are direct substitutes), the existence of a negative effect on the marginal productivity of private capital and employment could counterbalance the positive effects. This way, if increases in public capital imply negative effects on private capital (the direct substitution effects prevails for these regions), regional policy makers would have to implement measures in these regions that favor decreases in public capital. Among the negative domestic effects, Catalonia is the region with the lowest value; it is the most sensitive regional economy to a negative unit shock in public

Table 4: Regional Effects after 5 years

Regions	Domestic Effects	Mean Outward Effects (Weighted Average by GVA)	Mean Inward Effects (Weighted Average by GVA)	Mean Outward Effects / Domestic Effects	Mean Inward Effects / Domestic Effects
AND	-0,53	-0,18	-0,94	0,35	1,78
ARA	-0,38	-0,69	-0,11	1,82	0,28
AST	-0,44	-0,58	-0,10	1,32	0,24
BAL	-0,03	-0,38	-0,14	12,69	4,65
CAN	0,13	0,10	-0,15	0,77	-1,18
CANT	-0,99	-1,21	-0,02	1,22	0,02
CAT	-1,37	-1,30	-1,19	0,95	0,87
CLM	-0,37	-0,09	-0,29	0,25	0,79
CYL	-0,08	-0,19	-0,21	2,43	2,59
EXT	-0,05	0,44	-0,05	-8,80	1,02
GAL	-0,06	-0,16	-0,19	2,68	3,13
MAD	-0,70	-0,62	-1,16	0,89	1,65
MUR	-0,88	-0,86	-0,14	0,98	0,16
NAV	-0,44	-0,57	-0,10	1,29	0,22
PV	0,48	0,25	-0,50	0,52	-1,05
RIO	0,11	0,48	-0,01	4,32	-0,11
VAL	-0,17	-0,29	-0,56	1,72	3,28

Notes: Numbers reported are estimates of responses of GVA to one standard error negative shock to public capital in the region row or one standard error negative global shock to public capital.

capital.

With respect to the spillover effects shown in Table 3, it is also possible to find positive and negative effects. From the responses, it is important to highlight the existence of asymmetric shocks, ones that affect a pair of regions in a dissimilar way[11]. For example, a negative unit shock in the Valencian Community has negative effects on Andalusia, while a negative shock in Andalusia has a positive effect on Valencian Community. Another example is the case of the estimated effects between Catalonia and the Basque Country. For explanation, and in line with the applied methodology, it is not possible to rely on the nature of trade between these regions since the results are derived from a complicated process involving the tracing of paths of influence throughout the system.

Any attempt to analyze these spillovers will require the calculation of the average effects generated "from a region to the rest" (outward effects) and "from the rest to a region" (inward effects). The outward (inward) mean effects are computed from Table 3 by rows (columns) as weighted averages based on the regional shares of national GVA. The results are shown in Table 4. From columns 5 and 6 of this table, the estimation results show some evidence of the relative contribution of spillover effects on domestic effects. Ten of the Spanish regions (59%) have outward effects that are smaller than the corresponding domestic (internal) effects. In turn, for nine of the seventeen regions (53%), inward effects of public capital installed outside each region are more relevant than the domestic effects of public capital installed in the region.

The lowest negative outward effects can be found in Catalonia and Cantabria. Conversely, the highest positive effects (push-out growth spillovers) emanate from La Rioja, Extremadura and the Basque Country. Considering the range of values obtained by the domestic effects (X-axis) and the outward effects (Y-axis), different categories of regions can be depicted. Table 5 provides a taxonomy of the Spanish regional economies, highlighting the way in which reductions in public capital operate; different clusters of

[11] Chung, Hewings (2015) found similar complex typologies of interactions in an application in the Midwest state economies of the US.

Table 5: Regional taxonomy: Domestic Effects versus Mean Outward Effects

Mean Outward Effects	0.47		CLM	CAN, EXT, PV, RIO	
	-0.11				
			AND, ARA, AST, VAL, MAD, NAV	BA, CL, GA	
	-0.7				
	-1.3	CANT, CAT, MUR			
		-1.37	-0.75	-0.13	0.48
			Domestic Effects		

Table 6: Regional taxonomy: Domestic Effects versus Mean Inward Effects

Mean Inward Effects	-0.01	CANT, MUR	CLM, AST, NAV, ARA	CAN, EXT, RIO, GA, CL, BA	
	-0.40				
	-0.79		VAL	PV	
	-1.19	CAT	MAD, AND		
		-1.37	-0.75	-0.13	0.48
			Domestic Effects		

regions are observed And the classification of Table 5 generates some unexpected findings. For instance, negative shocks in the public capital of Cantabria, Catalonia and Murcia have more serious implications than others since they will produce important negative effects not only on their own regions but their corresponding trade-related regions. On the other hand, Canary Islands, Extremadura, the Basque Country and La Rioja are regions where the impact of austerity measures on public investment will produce lower negative effects than in the rest of regions.

Finally, Table 6 presents another taxonomy of the Spanish regional economies in view of the range of values obtained by the domestic effects (X-axis) and the inward effects (Y-axis). The different categories of regions recognize the importance of regional economic size in the generation of these effects, since all the regions, except the five largest regional economies (Catalonia, Madrid, Andalusia, Valencian Community and the Basque Country), have push-in effects above the mean. However, the larger regional economies are more sensitive to the negative shocks in public capital in the rest of regions.

The foregoing analysis suggests that a global reduction in the level of public capital stock will generate important implications for regional growth, and it is crucial, in designing regional growth-promoting strategies, to account for the variety of spatio-temporal effects through which regional public investment affect the rest of the regional economies.

5.1 Policy implications

From the economic literature, it would be expected that public investment in core regions would produce positive effects on regional growth due to the higher efficiency of investment in these type of regions (e.g. Caminal 2004), suggesting that regional public capital expenditure should be channeled to more developed regions (Blažek, Macešková 2010). For the Spanish case, De la Fuente (2004, p. 502) recommends investing "... a lot more in some of the richest regions and considerably less in some of the poorest ones." Nevertheless, the uneven distribution of public investment, favoring only developed regions, could produce negative results in terms of regional equity.

In view of the previous results, some policy implications can be suggested that will have implications for regional growth strategies. Taking account of the current public financial difficulties, the results suggest that decisions about the allocation of regional public investments can affect regional economic growth. The detected effects of public capital on regional growth indicate that some regions act as key generators of economic activity within the Spanish regional economic system.

From a policy standpoint, this new approach can provide some important insights. First at all, a negative global shock to public capital in Spain (first scenario) would not be a prudent development, since all the regions will suffer a decrease in regional growth. Secondly (second scenario), a key policy lesson also is that regional economic policy should not underestimate the benefits of maintaining levels of public capital in some regions: when the contribution to regional economic growth may be compromised in terms of system-wide effects (rather than just in a single region), it would be necessary to maintain public investments in these key regions. On the other hand, there are some options for reductions in public capital in the regions where public capital resources are going to be underused in terms of their contribution to regional economic growth. However, data on the effective utilization of existing capital stock are not available.

On the other hand, some important findings can be highlighted with respect to the reciprocal contributions between Catalonia and the rest of the Spanish regions. First, negative shocks in public capital of the rest of the Spanish regions produces negative effects (above the mean) in Catalonia. Thus, Catalonia is within the group of Spanish regions that are more sensitive to the negative shocks in public capital of the rest of the Spanish regions, and this indicates a higher dependence of Catalonia on the rest of Spain. Secondly and at the same time, negative shocks in the public capital of Catalonia have more serious implications than negative shocks in other regions. The conclusion is that our analysis of reductions in public capital in the Spanish regional economic system indicates that Catalonia occupies the first place among the most sensitive regional economies in Spain; Catalonia belongs to the group with lower spillover effects, and it is the region with the lowest inward effect.

6 Final remarks and conclusions

Public capital is shown to be a significant determinant of regional output growth. People, private capital and other regional factors are mobile across regions, and trade will connect regions, generating variations in the responses to public capital provision both in the short- and long-run. Nevertheless, neither of the dominant methods of analyzing the impact of public investments within a regional economic system considers the possibility that regional responses reflect in part the changes in public capital in other region(s). In this paper, an empirical framework is used that links the allocation of regional public investment with regional economic growth.

The aim of this research has been to implement a quantitative approach to guide and prioritize public investment within a closed regional economic system. The performance of the Spanish regions has been evaluated through the assessment of how output in every region within the country is affected by a reduction in the public capital stock of another specific region. This procedure can be considered as a simulation of the impact of austerity measures on public expenditure of the Spanish central and regional governments.

To accomplish this task, a multiregional integrated method has been used to investigate the regional growth effects of public capital in Spain. The specified MultiREG-SpVAR model facilitates addressing many of the econometric criticisms of the previous literature. In fact, the traditional VAR framework so often used in the empirical analysis excludes the presence of spatial feedbacks among the relevant variables of the regions.

The contribution to the literature is twofold. First, in contrast to existing studies that estimate 'local' VAR models, a 'global' VAR specification is used to test the effect of a reduction (although in other economic circumstances it could have been an increase) in the stock of public capital, accounting for heterogeneity, simultaneity and spatial autocorrelation. This econometric approach provides the opportunity to identify both domestic and spillover effects of a reduction in public capital on regional output growth. Secondly, the analysis moves beyond most existing cross-section studies by providing more information about the location where the spillover effects originate. These contributions allow identification of regions in which further public capital reduction (increase) would have a lower (greater) effect on the output.

In particular, the responses of regional outputs to a negative unit shock in regional public capital have been analyzed using two different scenarios: a) a negative global shock to public capital in Spain; and, b) a negative unit shock to public capital in each region. From the first scenario, the detected effects appear quite plausible; the results show that reduced funding for public capital accumulation will negatively affect the economic growth of all the regions. The conclusion is that global decreases in public investment have a homogeneously negative effect on the output of all the regions. From the second scenario, the domestic and the spillover effects display positive and negative signs. Although negative effects prevail, the existence of positive effects can guide regional public capital allocation, prioritizing regional investment.

The empirical analysis reveals the existence of different clusters of regions. From this taxonomy, some regions facing substantial cuts in public investment may be able to maintain a modest level of regional growth over time. On the other hand, the diminution of public capital in other regions will negatively affect the regional economic system. In other words, the Spanish regions portray heterogeneous responses from localized public capital reductions over the years considered. This means that there is a place for strategic, planning actions from policy-makers in Spain to improve the allocation of public resources: it may be possible to partially mitigate the impact of austerity measures on government spending through a more considered, spatially targeted strategy. However, such a strategy presents significant political challenges to a central government interested in balancing concerns with national growth and development with attention to regional equity.

Although the findings in this paper make an important contribution to the ongoing debate about which regional policies need to be promoted to raise and sustain Spanish regional economic growth, some cautionary remarks seems appropriate. First, the analysis was conducted at the regional level; it would be interesting to considerer a multilevel analysis, since, at a lower level of disaggregation, regional responses may be even more heterogeneous and with a concomitant increase in spillovers. Disaggregation of public capital into its main components might also reveal some greater variations. At the same time, some integration of this modeling approach inside a multiregional general equilibrium model might enhance the understanding of the full system-wide impacts, especially the ways in which reductions in public investment affect employment levels, consumption, production, and government revenues.

Acknowledgement

We thank the editor and two anonymous referees for their comments. The authors acknowledge financial support from the Spanish Ministry of Economy, Industry and Competitiveness (Grant ECO2016-78352-P).

References

Agénor PR, Moreno-Dodson B (2006) Public infrastructure and growth: New channels and policy implications. Policy research, working paper, no. 4064, world bank. CrossRef.

Álvarez A, Arias C, Orea L (2006) Econometric testing of spatial productivity spillovers from public capital. *Hacienda Pública Española* 178[3]: 9–21

Aschauer DA (1989) Is public expenditure productive? *Journal of Monetary Economics* 23[2]: 177–200. CrossRef.

Baxter M, King RG (1993) Fiscal policy in general equilibrium. *The American Economic Review* 83[3]: 315–334

Blažek J, Macešková M (2010) Regional analysis of public capital expenditure: To which regions is public capital expenditure channelled – to 'rich' or to 'poor' ones? *Regional Studies* 44[6]: 679–696. CrossRef.

Boarnet MG (1998) Spillovers and locational effects of public infrastructure. *Journal of Regional Science* 38: 381–400. CrossRef.

Bustos A, Cutanda A, Díaz A, Escribá FJ, Murgui MJ, Sanz MJ (2008) La BD.MORES en base 2000: nuevas estimaciones y variables. Working paper d-2008-02, dirección general de presupuestos, ministerio de economía y hacienda

Caminal R (2004) Personal redistribution and the regional allocation of public investment. *Regional Science and Urban Economics* 34[1]: 55–69. CrossRef.

Capello R (2009) Spatial spillovers and regional growth: a cognitive approach. *European Planning Studies* 17: 639–658. CrossRef.

Cheshire P, Carbonaro G (1996) Urban economic growth in Europe. *Urban Studies* 33: 1111–1128

Chung S, Hewings GJD (2015) Competitive and complementary relationship between regional economies: A study of the Great Lake States. *Spatial Economic Analysis* 10: 205–229. CrossRef.

Cohen J, Morrison Paul C (2007) The impacts of transportation infrastructure on property values: A higher-ordered spatial econometrics approach. *Journal of Regional Science* 47[3]: 457–478. CrossRef.

De la Fuente A (2004) Second-best redistribution through public investment: a characterization, an empirical test and an application to the case of Spain. *Regional Science and Urban Economics* 34: 489–503. CrossRef.

De la Fuente A (2010) Infrastructures and productivity: an updated survey. UFAE and IAE working paper 831.10, Unitat de Fonaments de l'Anàlisi Econòmica (UAB) and Institut d'Anàlisi Econòmica (CSIC)

Dees S, Di Mauro F, Pesaran MH, Smith LV (2007) Exploring the international linkages of the euro area: a global VAR analysis. *Journal of Applied Econometrics* 22: 1–38. CrossRef.

Deliktas E, Önder A, Karadag M (2009) The spillover effects of public capital on the Turkish private manufacturing industries in the geographical regions. *The Annals of Regional Science* 43[2]: 653–378. CrossRef.

Deno K (1988) The effects of public capital on U.S. manufacturing activity: 1970 to 1987. *Southern Economic Journal* 53: 400–411. CrossRef.

Destefanis S, Sena V (2005) Public capital and total factor productivity: New evidence from the Italian regions 1970-98. *Regional Studies* 39[5]: 603–617. CrossRef.

Devarajan S, Swaroop V, Zou HF (1996) The composition of public expenditure and economic growth. *Journal of Monetary Economics* 37: 313–344. CrossRef.

Dietzenbacher E, van der Linden JA, Steenge AE (1993) The regional extraction method: EC input-output comparisons. *Economic Systems Research* 5: 185–206. CrossRef.

Garcia-Mila T, McGuire TJ, Porter RH (1996) The effect of public capital in state-level production functions reconsidered. *The Review of Economics and Statistics* 78[1]: 177–180. CrossRef.

Gramlich EM (1994) Infrastructure investment: a review essay. *Journal of Economic Literature* 32[3]: 1176–1196

Heintz J (2010) The impact of public capital on the US private economy: new evidence and analysis. *International Review of Applied Economics* 24[5]: 619–632. CrossRef.

Henderson D, Kumbhakar S (2006) Public and private capital productivity puzzle: A nonparametric approach. *Southern Economic Journal* 73[1]: 219–232. CrossRef.

Holtz-Eakin D (1994) Public-sector capital and the productivity puzzle. *The Review of Economics and Statistics* 76[1]: 12–21. CrossRef.

Holtz-Eakin D, Schwartz A (1995) Spatial productivity spillovers from public infrastructure: Evidence from state highways. *International Tax and Public Finance* 2: 459–468. CrossRef.

Hunt C (2012) The interaction of public and private capital: a study of 20 OECD members. 44: 739–764. CrossRef.

Kamps C (2005) The dynamic effects of public capital: VAR evidence for 22 OECD countries. *International Tax and Public Finance* 12[4]: 533–558. CrossRef.

Koop G, Pesaran MH, Potter SM (1996) Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics* 74: 119–147. CrossRef.

Llano C, Esteban A, Pérez J, Pulido A (2008) La base de datos C-intereg sobre el comercio interregional de bienes en España: método y primeros resultados (1995-2006). *Ekonomiaz* 69[3]: 245–269

Llano C, Esteban A, Pérez J, Pulido A (2010) Opening the interregional trade black box: The C-Intereg database for the Spanish economy (1995-2005). *International Regional Science Review* 33[3]: 302–337. CrossRef.

Lynde C, Richmond J (1992) The role of public capital in production. *Review of Economics and Statistics* 74: 37–44. CrossRef.

Mas M, Pérez F, Uriel E (2009) El stock y los servicios del capital en España y su distribución territorial. Fundación BBVA-Ivie

Mas M, Pérez F, Uriel E (2011) El stock y los servicios del capital en españa y su distribución territorial y sectorial (1964-2010). Technical report

McGregor PG, Swales JK (2003) The economics of devolution/decentralisation in the UK: Some questions and answers. Discussion paper in economics 03-13, University of Strathclyde, Glasgow (http://www.strath.ac.uk/media/departments/economics/researchdiscussionpapers/2003/media_34453_en.pdf)

Melitz MJ, Ottaviano GIP (2008) Market size, trade, and productivity. *Review of Economic Studies* 75: 295–316. CrossRef.

Moreno R, López-Bazo E (2007) Returns to local and transport infrastructure under regional spillovers. *International Regional Science Review* 30: 47–71. CrossRef.

Morrison P, Schwartz A (1996) State infrastructure and productive performance. *The American Economic Review* 86[5]: 1095–1111

Márquez MA, Ramajo J, Hewings GJD (2010a) Measuring the spillover effects of public capital: A bi-regional structural vector autoregressive analysis. *Letters in Spatial and Resource Sciences* 3: 111–125. CrossRef.

Márquez MA, Ramajo J, Hewings GJD (2010b) A spatio-temporal econometric model of regional growth in Spain. *Journal of Geographical Systems* 12: 207–226. CrossRef.

Márquez MA, Ramajo J, Hewings GJD (2011) Public capital and regional economic growth: a SVAR approach for the Spanish regions. *Investigaciones Regionales* 21: 199–223

Márquez MA, Ramajo J, Hewings GJD (2015) Regional growth and spatial spillovers: Evidence from an SpVAR for the Spanish regions. *Papers in Regional Science* 94: S1–S18. CrossRef.

Munnell A (1990) Why has productivity declined? Productivity and private investment. New England Economic Review, Federal Reserve Bank of Boston, January-February (https://www.bostonfed.org/-/media/Documents/neer/neer190a.pdf)

Pereira AM, Roca-Sagalés O (2003) Spillover effects of public capital formation: Evidence from the Spanish regions. *Journal of Urban Economics* 53: 238–256. CrossRef.

Pesaran MH, Schuermann T, Weiner SM (2004) Modelling regional interdependencies using a global error-correcting macroeconometric model. *Journal of Business and Economic Statistics* 22[2]: 129–162. CrossRef.

Pesaran MH, Shin Y (1998) Generalized impulse response analysis in linear multivariate models. *Economic Letters* 58: 17–29. CrossRef.

Ramajo J, Márquez MA, Hewings GJD (2017) Spatiotemporal analysis of regional systems: A multiregional spatial vector autoregressive model for Spain. *International Regional Science Review* 40: 75–96. CrossRef.

Ratner JB (1983) Government capital and the production unction for U.S. private output. *Economics Letters* 13: 213–217. CrossRef.

Roca-Sagalés O, Sala H (2006) Efectos Desbordamiento de la Inversión en Infraestructuras en las Regiones Españolas. *Investigaciones Regionales* 8: 143–161

Romp W, De Haan J (2007) Public capital and economic growth: A critical survey. *Perspektiven der Wirtschaftspolitik* 8: 6–52. CrossRef.

Seitz H (1993) A dual economic analysis of the benefits of the public road network. *The Annals of Regional Science* 27[3]: 223–239. CrossRef.

Seitz H, Licht G (1995) The impact of public infrastructure capital on regional manufacturing production cost. *Regional Studies* 29: 231–240. CrossRef.

Serrano Martínez L, Mas Ivars M, Pérez García F, Uriel Jiménez E (2017) Acumulación y productividad del capital en España y sus comunidades autónomas en el Siglo XXI. Fundación BBVA, Bilbao

Sims CA (1980) Macroeconomics and reality. *Econometrica* 48[1]: 1–48. CrossRef.

Sturm JE, Kuper GH, de Haan J (1998) Modelling government investment and economic growth on a macro level: a review. In: Brakman S, van Ees H, Kuipers SK (eds), *Market Behaviour and Macroeconomic Modelling*. MacMillan Press Ltd, London, 359–406. CrossRef.

Tatom J (1991) Public capital and private sector performance. *Federal Reserve Bank of St.Louis Review* 73: 3–15

Young A (1991) Learning by doing and the dynamic effects of international trade. *Quarterly Journal of Economics* 106: 369–405. CrossRef.

Yu N, de Jong M, Storm S, Mi J (2013) Spatial spillovers of transport infrastructure: evidence from Chinese regions. *Journal of Transport Geography* 28: 56–66. CrossRef.

0 0 1 2 6 2 1 4 1 2

Funded by

ersa

WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

FWF