DOES CBD THEORY SURVIVE THE TEST OF SMALL CITIES? CITY-SIZE AND SPRAWL IN ITALY

Gianni GUASTELLA
Università Cattolica, Dep of Mathematics and Physics, 41Via Musei, 25121, Brescia (IT)
giovanni.guastella@unicatt.it
Corresponding author

Stefano PAREGLIO
Università Cattolica, Dep of Mathematics and Physics, 41Via Musei, 25121, Brescia (IT)
stefano.pareglio@unicatt.it

Abstract

Economic theory predicts that the equilibrium of different economic forces explains the spatial scale of a city more than the uncontrolled take of agricultural land, which is considered instead as urban sprawl. A wide range of empirical results based on US data for large urban areas supports this hypothesis, showing that the socio-economic and environmental forces explain a vast portion of the variation in urbanization across cities. In this paper, we ask whether these socio-economic forces are relevant also in small cities and if they are in a different manner, provided that sprawling phenomena may occur more easily in small areas due to the larger availability of agricultural land. To answer the question, we estimate the relationship between city size and the socio-economic and environmental forces using data for small and large municipalities in the Lombardy region, Italy, and test to what extent this model is apt to explain size variations. We find that the model is adequate also in the case of small cities but differentiating small from large cities suggests that the sprawl hypothesis cannot be ruled out by the empirical evidence as the process of land conversion from agricultural to urban is substantially faster in small and medium-sized cities compared to large ones.

Keywords: Land Use, Urban Sprawl, Central Business District, Spatial Econometrics, Italy

JEL classification: O18, Q15, R14

1. Introduction

The Alonso-Mills-Muth (AMM) model of the urban spatial structure is used in urban economics literature to study the size and structure of urban agglomerations. The pioneering study of Brueckner and Fansler [5] has set up the empirical framework to investigate the extent to which the AMM predictions are apt to explain variation is the spatial size of cities. Urban size is predicted to increase with increasing income and population and to decrease with increasing transport costs and agricultural rents. More recent studies (McGrath [11]; Paulsen [14]; Spivey [16]) contributed to assessing the validity of AMM predictions empirically. The evidence clearly supports this theoretical model, proving that the monocentric development of cities built on the balance between the socio-economic pressures for urbanization and the agricultural and environmental constraints to land use change continues to be the reference model of urban expansion.

Understanding the determinants of the spatial scale of cities is even more important in present times. The urban spatial expansion is frequently associated with the risk of sprawl, which denotes an excessive waterproofing of land, coining the possibility that urban growth subtracts production factors to agriculture, altering the equilibrium of natural resources' use (Brueckner and Fansler [5]). Consequently, national and local policy makers are called for the treatment of the phenomenon by limiting urban spatial expansion and regulating land use change (Brueckner [4]). The AMM model, linking the urban size to its main economic determinants, can be used to explain the extent of an urban size that is determined by the economic market forces, as predicted by the model. Significant coefficient estimates confirm that urban expansion relates to socio-economic rationales more than to the unregulated take of agricultural land by urban settlements.

However, it remains debatable that the empirical model initially proposed by Brueckner and Fansler [5] is apt to address the incidence of urban sprawl in light of the more recent urban expansion dynamics. Since the test builds on a simple relationship between the size of urbanized area and socio-economic variables, it is not clear what magnitude of these effects would allow excluding the hypothesis of urban sprawl. For instance, the evidence that growth follows demographic trends in almost all cities it is not sufficient to rule out the sprawl hypothesis. In contrast, it is a matter of fact that the consumption of land is excessive in certain cities, at least compared to what can be expected based on the economic and demographic figures. Accordingly similar changes in population may produce very different patterns of land use change across cities; in contrast the simple regression approach produces an average estimate that may fail to capture such heterogeneity. Furthermore, the estimates provided by the literature do not present a single benchmark of what could be considered an acceptable response, measured as the change in land use, to the change in the socio-economic determinants. Perhaps this is because different methodological approaches have been used to retrieve coefficient estimates, or because of the differences in the estimation samples. In fact, this literature has been concerned primarily with cities in the US (Brueckner and Fansler [5]; McGrath [11]; Paulsen [14]; Spivey [16]). Except Deng et al. [8] and Song et al. [15], which focused on China, and Brueckner and Sridhar [6], which focused instead on India. Furthermore, in all these studies evidence is provided based on data of large urban agglomerations only.

The present paper concerns the investigation of the AMM predictions in a European territory, namely the Lombardy region, in Italy, among the most densely populated regions in the country. Building on this consolidated empirical literature, the research in this paper extends the geographical scope of the analysis by testing the hypothesis upon data of all the cities in the region and not of large urban agglomerations only. Regarding the contribution to the empirical literature, this paper is aimed at investigating the extent to which the AMM predictions may also apply to small and medium-sized cities. While AMM predictions are not expected to be relevant for large cities only, at least on a theoretical ground, the analysis of medium-sized urban areas has been likely prevented the lack of appropriate data on urbanized area at small geographical scales. Notwithstanding, sprawl is becoming an important phenomenon in medium-sized cities, where speculative behaviors that leverage on the availability of natural and agricultural land may find fertile soil, more than in large cities, where market forces are clearly dominant instead. Such a heterogeneous sample allows further conducting the research by exploring the varying incidence of both market forces and sprawl across cities of different scales. Following a two-step empirical strategy, the model in Brueckner and Fansler [5] is estimated for all cities in the regions first and then allowing for structural instability across groups of cities in the model' intercept and slopes.

The empirical analysis builds upon a unique dataset set up by the Lombardy region that includes information, for the year 2007, on soil destination at the municipality level (there are 1568 municipalities in the estimation sample). This dataset allows determining, for each municipality, the city size as measured by the number of square km of urbanized area. The use of data on contiguous municipalities requires using appropriate spatial econometric techniques to consider spatial relations in the estimation of model equations.

Econometric results suggest that the AMM model is apt to explain city size also in the case of small and medium-sized urban areas, but evidenced some specific issues in the results. Firstly, an unexpected negative income effect occurs as a consequence of the inclusion of contiguous areas in the estimation sample. In the AMM model consumers are not allowed to choice the residential location across different cities as the only distance from the Central Business District (CBD) is considered in the utility-maximization problem. In contrast, cross-cities house-to-work commuting is quite a common phenomenon in recent times and individuals working in a CBD may prefer to live in neighboring cities where, commuting time being almost constant, the housing good is less expensive. While such an effect might be negligible in the comparison across large urban agglomeration, which are usually placed at long distances, it might be substantially relevant to the comparison of neighboring municipalities within a region. Secondly, the observed agricultural prices may poorly proxy agricultural rents, as actual market prices already discount the value of future agricultural land reconversion for urbanization purposes. Such an effect is especially relevant in large cities.

Finally, the influence of market forces on urban size is substantially lower in the case of small and medium-sized cities, in which sprawl occurs more frequently. In Lombardy the average urbanized square meter per inhabitant has grown by 0.06 during the period 1999-2007 in cities with more than 20000 inhabitants, varying from 2.55 to 2.61. The same figure rises to the value of 0.247 in the case of municipalities with less than 5000 inhabitants, in which the average urbanized square meter per inhabitant has grown from 6.023 1999 to 6.27 in 2007. These descriptive statistics characterize small cities for the substantially larger values of both use and take of land. In marginal terms, the same result is confirmed by the econometric evidence.

The remainder of this paper is structured as follows. The next section surveys empirical literature about the AMM model and discusses the empirical issues related to empirical estimation based on data on units of small geographical scale. Section three introduces the dataset and the empirical model. Section four summarizes the results. A discussion concludes the work.

2. The empirics of city size distribution

City size distribution and urban growth are key themes in the urban economics literature. While the mono-centric organization of urban space has prevailed until the first half of the 20th century, urban decentralization, scatteration, and sprawl have characterized the urban development of increasingly polycentric cities in the last decades (Glaeser and Kahn [9]). Sprawl and urban growth appear as two faces of the same medal. Urban agglomeration generates higher income and attracts more workers increasing, in turn, the demand for housing and land. As consequence, the urban fringe expands toward peri-urban and rural spaces, causing negative economic and environmental externalities that markets usually fail to take into account, and, for this reason, sprawl is frequently related to inadequate urban planning policies. Sometimes urbanized areas expand to an extent that is larger than what it could be reasonably expected based on the agglomeration of people and firms in the cities, resulting in the loss of agricultural land, longer commuting, and ultimately low urbanization densities. In this respect, sprawl is characterized by an excessive urban expansion (Brueckner [4]).

Cities are the engine of economic growth, which comes as a consequence of urbanization externalities driven either by industrial specialization in small and medium size cities or by industrial variety in large metropolitan areas. Likewise, the socio-economic dynamics are responsible for the growth of cities which happens at the expenditure of the agricultural sector and, in general, of the environment. As urbanization pressures increase, the growth of cities might threaten the ecological equilibrium of the territories and hamper their potential for rural development, with dramatic social consequences such as massive migration from the countryside and also dramatic ecological consequences that include an increased risk of floods and the deterioration and agricultural land. For this reason, it has now become essential for policy makers to understand the relationship between city size and market, in a way to determine how far the city is from the optimal size and eventually which policies are required to regulate and definitively curb an excessive consumption of land. With a tradition originating in the so-called Alonso-Mills-Muth (AMM) model, urban economist attempted to study this relationship by explaining urban expansion as a function of income, population, transport costs, and agricultural land. Despite the restrictive theoretical assumptions about homogeneous incomes and preferences, some empirical papers provided robust evidence in support of the AMM model (Brueckner and Fansler [5]; McGrath [11]; Paulsen [14]; Spivey [16]; Wassmer [17]). In summary, it is confirmed that the growth of urbanized land in cities is substantially determined by market forces rather than being the result of an uncontrolled consumption of agricultural land.

Under the hypothesis of the AMM model, the urban fringe defines the optimal size of the monocentric city and is located at the critical distance from the Central Business District (CBD) where the housing rent equals at least the agricultural one, and the level of utility is the same for all the households. Some households, in fact, prefer a larger house, in the periphery, and hence accept a longer commuting to work. McMillen [12] classifies the empirical approaches to test the AMM hypothesis in two broad categories. One category includes the regression-based approaches. A log-level equation is estimated where housing prices, land values, capital-land ratios, or population densities are the dependent variables, and the distance from CDB explains their spatial variation. A negative and significant slope (density gradient) is confirmatory evidence of the hypothesis mentioned above. The other category includes models based on the comparative statics initially formulated by Wheaton [18] and summarized here:

$$\frac{\partial A}{\partial P} > 0; \frac{\partial A}{\partial Y} > 0; \frac{\partial A}{\partial R} < 0; \frac{\partial A}{\partial T} < 0. \qquad \textbf{(1)}$$

In the equation (1), A is the urbanized area of a city, Y is the median households' income, R is the agricultural land rent and, finally, T is the unitary measure of transportation costs. Since AMM model represents the city as a closed world, comparative statics are especially expected to hold when observing a single city over time. Based on McMillen [12],predictions are testable also using cross-sectional information on different cities. In this latter case a simple linear regression model is specified, for a series of $i = 1, 2, ..., N$ cross-sectional units as in equation (2).

$$A_i = \beta_0 + \beta_1 \cdot P_i + \beta_2 Y_i \cdot + \beta_3 R_i \cdot + \beta_4 \cdot T_i + e_i \qquad \textbf{(2)}$$

Because an increase in population will shift the demand for housing, more houses will be demanded by individuals at the edge of the city and hence the expected value of the coefficient β_1 is larger than zero. An increase in income likely produces similar effects, rising the demand for larger houses and space; that is available at lower prices at the edge of the city. Nonetheless, when income increases the individuals' aversion to commuting also increases, making the demand for housing higher at the CBD, with possible negative consequences for the city size (McMillen [12]). An additional negative effect of income on size may derive from cross-cities commuting. While the city is a closed world in the traditional AMM structure, individuals, in fact, have the option to live also outside the edge of the city where they work and to commute between cities. The urban studies literature has documented trends in urban decentralization (Cervero and Wu [8]), specifically in large cities, where both income and house prices are substantially higher. As individuals may prefer to buy houses at a lower price in the neighboring cities, small centers in the proximity of large urban agglomeration may exhibit relatively higher income levels leading to a negative relationship between income and urban size. Accordingly, the coefficient β_2 might be either positive or negative depending on which effect dominates: although the empirical studies that analyze diverse cities frequently evidenced the case for a positive relationship, a negative one is more likely observed in the case of contiguous territorial units located in the same region.

City expansion occurs in AMM model extending the radius of the urban fringe, converting agricultural territories into urbanized areas. A high productivity in the agricultural sector makes farmland more expensive and, other things being equal, the housing market clears at a lower distance from the CBD. For this reason, the relationship between urban size and the agricultural rent reflected in the associated slope coefficient (β_3) is expected negative.

Finally, it is more convenient for individuals to live near the CDB when the unitary commuting costs increase; that means that the city size is relatively smaller small when the transport costs are high, causing the expected value of the coefficient β_4 to be negative.

A number of papers, mostly concerned with the US, provided cross-sectional evidence related to the empirical model in the equation (2). Brueckner and Fansler [5] analyze data of 40 urban areas of medium size, with populations ranging between 52,000 and 257,000 inhabitants, in 1970. The study evidences a clear consistency between theory and empirics, being almost 80% of cross-sectional variation in urban area explained by the model's predictors. Coefficient estimates related to all variables but the transport costs all are correctly sloped and significant. McGrath [11] updates the study by Brueckner and Fansler [5] by using longitudinal-data of the 33 largest US metropolitan areas (populations range between 136,000 and 16,207,000 inhabitants) for five decades (1959-1990). Empirical results are confirmatory of the early evidence in Brueckner and Fansler [5]. The transport cost variable is statistically significant and the total variation explained by the model is about 90%. Therefore, one more indication is given that urban growth in US cities is the result of market forces to the greatest extent, ruling out the hypothesis of an uncontrolled take of agricultural land. Wassmer [17] uses this empirical framework to test the effectiveness of local urban containment policies. The model estimates a cross-section of 452 US urban areas with a population ranging between 50,058 and 17,799,681 inhabitants. Both the population and the income effects are positive and significant, as suggested by the theory while the effect of the agricultural land prices is insignificant. The study does not consider te effect of transportation costs. In Spivey [16] the same sample of 452 cities used in Wassmer [17] is analyzed alongside an additional sub-sample made of 85 large cities only. For both samples the relationship of the urban area with population is positive; the relationship with income is negative in the enlarged sample, and it turns positive in the sub-sample of large cities. Finally Paulsen [14], similar to McGrath [11], uses panel data methods to analyze the urbanization of 329 cities across three decades (1980-2000). For both population and income, the relationship with urbanized area is positive and statistically. Also this study excludes the effect of transportation costs.

All previous studies estimate parameters in the β vector in equation (2) using either cross section or panel data methods and assuming independence between $i = 1, 2, ..., N$ cross-sectional units. The assumption of independence between the components of the error vector (e_i) in equation (2) may be unreasonable, however, when the cross-section units are nearby-located in space. In the alternative, a spatial error structure (Anselin [1]) of residuals can best account for the geographical relations occurring among cities. More in general, some reasons motivate the use of spatial econometrics in models for contiguous geographical areas, stemming from the unobserved spatial heterogeneity to the omission of spatially correlated variables from the model specification to the presence of contagion effects in the data generating process (LeSage and Pace [10]). In models for the urban spatial scale, these motivations translate in the formulation of two relevant hypothesis concerned with spatial relations between geographical areas. On the one hand, individuals can make their residential choice among neighboring cities, not only among different locations at varying distances from the CBD. That intrinsically extends the trade-off between costs and benefits of commuting to the case of more cities, coining the possibility that an increase in the demand for urbanized area in a city also has effects on its neighbors. In this case, the spatial relations operate through the dependent variable. On the other hand, the size of a city's urban area is conditioned by external factors specific to that city. Being these factors usually unobservable to the econometrician, they contribute to form the vector of residuals. If these factors are also unevenly distributed in the geographical space of the sample, the spatial concentration moves to the disturbances causing their spatial correlation. In this case, spatial relations operate through the error term. In the present study, Lagrange Multiplier diagnostics (Anselin [2]) are used to discriminate between the so-called "Spatial Lag Model" (LAG), that includes spatial relations in the dependent variable, and the "Spatial Error Model", that includes spatial effects in the error term.

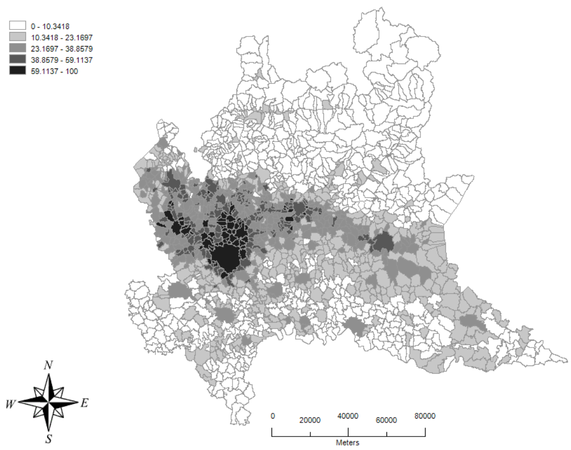
In addition to spatial relation across units, the study of contiguous geographical areas requires addressing a second issue, concerned with income. The income effect is expected positive when several independent spatial units are compared in the estimation. In contrast, such an effect may turn negative if the sample is made of contiguous geographical areas, since commuters' mobility more likely extends beyond the city's administrative borders. In the standard AMM model workers commute between the CBD and the urban fringe at a constant unitary commuting cost, such that total cost depends uniquely on commuting distance. In modern cities, networks of infrastructures facilitate the connections between cities making the cost of commuting between nearby cities relatively (and possibly absolutely) lower than the cost of commuting from the fringe to the CBD. Because individuals who work in a city trade-off between the house good and the costs of commuting, an increase in housing price following an increase in the income of the city may induce workers to buy a house in neighboring cities. In some circumstances, low inter-city commuting costs are also addressed as a cause of sub-urbanization and the emergence of secondary sub-centers (Cervero and Wu [7]). In the context of empirical models for urbanized area, evidence of a negative income effect has been already reported by Spivey [16].

An ultimate concern arises about spatial heterogeneity. By comparing urban units of very different scales, the structural stability of coefficient estimates can be doubted and, in contrast, dissimilarities in the effects across urban units of different size can be hypothesized. More specifically, income and population elasticities are expected larger in small urban areas, as more space is available for conversion in these territories and local administrations urge less against unnecessary urban expansion. On the contrary, there is uncertainty, about the nature of the difference between cities concerning agricultural rents and transport costs. Because the model has been usually estimated for large cities there is little empirical evidence available. To the authors' knowledge the only attempt to analyze the variation in effects across cities of different sizes has been made by Paulsen [14] finding that, based on the sample used for estimation, coefficient estimates statistically differ between large (>500000 inhabitants) and small cities.

3. Data and Model

The dependent variable used in the regression model is the size of the urbanized area in a municipality, measured in square km. The variable is collected by the Lombardy Region at irregular time intervals and is part of a larger project aimed at the construction and maintenance of an Agricultural and Forestry Soil Use Database (DUSAF). Data is available to the public directly from the institution upon request. Figure 1 shows the spatial distribution of the variable in the territory of the Lombardy region (year 2007).

Figure 1: Urbanization density (urbanized over total area) in Lombardy, 2007



The population (P_i) is measured as the total number of inhabitants in the municipality in the year 2007 and is obtained from the national institute of statistics (ISTAT). We measure the population over the total (the sum of urbanized, agricultural and forestry) area, to correct this measure for differences in population caused by the varying size of administrative units (municipalities).

Income in the municipality (Y_i) is the average income per inhabitant in year 2007. The information is part of the inquiry on income in municipalities carried out by "Il Sole 24 Ore", the most important business newspaper in Italy, and derives from the elaboration of data on fiscal contribution collected by the Ministry of Economics and Finance.

The value of agricultural rent (R_i) is from the database of land values maintained by the national institute of agricultural economics (INEA). INEA collects data on the values of agricultural land by type of farming activity at the province (NUTS III in Eurostat classification) level by altitude zone. For each province, an aggregate measure of land value is constructed as the weighted average of farm type specific values and using farm type area shares of Utilized Agricultural Area (UAA) as weight. This value is matched with information on the altitude zone of the municipality to retrieve the average value of land at the municipality level. Because land values vary only among provinces and, within each province among altitude zones, by construction, there is no variation in data across municipalities in the same province and at the same altitude zone. Notwithstanding, the measure is deemed representative of the differences in land markets across municipalities in the region, and it is by far the most detailed information that is possible to retrieve at such a specific level of territorial disaggregation.

An index of car use proxy the transport costs $\left(T_i\right)$. Following a consolidated literature, it is assumed that cars reduce transport costs significantly (Glaeser and Kahn [9]).Other things being equal, low congestion encourages the use of private transportation, reducing the cost of commuting for individuals. Information about the number of circulating vehicles $\left(c_i\right)$ is provided by the Italian automobile club (ACI) at the municipality level for the year 2007. Because the variable varies to a significant extent across municipalities, it is standardized by the sample average and weighted by the radial distance separating the CBD and the administrative border of the municipality. Consequently, a higher value of T_i is associated with a higher cost of commuting.

$$T_i = \left(\frac{c_i}{\sum_i c_i} \right)^{-d_i} \quad \textbf{(3)}$$

The original linear relationship in equation (2) employed to test AMM predictions is extended to consider externalities between municipalities either in the dependent variable (equation(4)) or in the error term (equation (5)). $X = \left[1, P, Y, R, T\right]$ is the matrix of model covariates and β is the vector of parameters to be estimated. LM tests (Anselin [2]) on linear model residuals guide the choice of the most appropriate model. W is the N-dimensional square matrix, N being the number of municipalities in the sample, incorporating necessary information on contiguity between geographical units. Inverse squared distance is used to weight contiguity relationships and, as usual, the matrix is row-standardized to obtain, when pre-multiplied by a vector, the average value of the vector in the neighbors of the region i.

$$A_i = \rho W A_i + X_i{}'\beta + e_i \quad \textbf{(4)}$$

$$\begin{cases} A_i = X_i{}'\beta + e_i \\ e_i = \lambda W e_i + u_i \end{cases} \quad \textbf{(5)}$$

The significance of coefficients in the β vector assess the validity of the AMM hypothesis. Four regimes are defined, and regime-specific coefficients are estimated to test for structural stability of coefficients in the β vector and to explore the extent of validity of AMM hypothesis in small and medium-sized cities, not only in large urban agglomerations. Regimes are defined based total population in the municipality, being this measure the most frequently employed in defining city size. As in Paulsen [15], structural stability is examined by using Chow tests and, more in detail, an alternative version of the test that is apt to analyze structural instability in spatial regression (Anselin [3]).

4. Results

Table 1 summarizes the first set of results. An attempt is made to assess the validity of AMM hypothesis estimating the linear model in equation (2) and including population and income variables only. In column (a) the estimated coefficient for the population variable is positive and statistically different from zero confirming, as expected, the robustness of the relationship. The estimated coefficient for income is, oppositely, negative and significant. Such evidence of a negative income effect is related to the use of small and contiguous geographical areas for estimation. As a result, cities with higher income levels offer houses at a relatively higher price and, accordingly, people may prefer moving to small and medium-sized cities, possibly in the neighborhood of the large urban agglomerations.

The model in column (b) additionally includes agricultural rent. In this case also, the slope of the coefficient, exhibits opposite to what predicted by theory. The estimate is, in fact, positive and largely significant. Nonetheless, in modern urban agglomerations, it is likely that agricultural land prices are determined by urbanization more than the opposite, ant this reflects in a positive relationship between land rents and the urban size, a relationship that is more likely to hold in the case that contiguous urban areas are analyzed. Contiguous territories are in fact more homogeneous regarding their structural characteristics affecting agricultural productivity and, consequently, price differences may be explained by factors other than those related to agriculture. In particular, among the several determinants, urbanization pressures are among the most important ones. In contrast, when comparing heterogeneous urban agglomerations, variation in land values might be associated with characteristics influencing agricultural productivity more than urbanization pressures, which are somehow homogeneous across urbanized cities. Finally, this evidence might be strongly associated, perhaps in the case of this paper, to the use of a proxy variable for land rents with relatively small territorial variation.

Attempting to narrow the specification problems caused by the potential endogeneity of the agricultural rent, in the absence of a valid instrument to overcome the problem completely, the variable is excluded from the estimation, possibly at the price of under-specification. Estimation results are presented in the column (c).The coefficient related to income turned now insignificantly different from zero, and there is no relevant change associated with the coefficient for the population. Finally, there is evidence of a negative effect of transportation costs, as predicted by AMM.

The regression model in column (d) includes all variables are. It is confirmed the validity of the AMM model of urban spatial structure through coefficients relate to population and transport costs. In contrast, income and agricultural rents coefficients show a sign opposite to the ones predicted by the theoretical model. All coefficients are significant and, overall, almost 90% of the total variance in urban size of municipalities is explained by covariates suggested by AMM.

Table 1: Test of CBD hypothesis in Lombardy – 2007

	(a)	(b)	(c)	(d)
Intercept	161.820***	99.099***	293.800***	206.000***
	(16.801)	(16.459)	(24.150)	(23.830)
Y	-23.803***	-45.752***	0.469	-25.500***
	(9.086)	(8.706)	(9.498)	(9.215)
P	0.117***	0.116***	0.134***	0.130***
	(0.001)	(0.001)	(0.003)	(0.002)
R		3.572***		3.371***
		(0.256)		(0.255)
TC			-653.800***	-512.400***
			(87.330)	(83.470)
Adj R2	0.864	0.879	0.868	0.882
Moran's I	0.269	0.189	0.240	0.170
	[0.00]	[0.00]	[0.00]	[0.00]
LM LAG	420.48	178.75	297.64	120.82
	[0.00]	[0.00]	[0.00]	[0.00]
LM ERR	990.24	486.26	787.77	393.64
	[0.00]	[0.00]	[0.00]	[0.00]
RLM LAG	27.50	7.93	13.55	1.87
	[0.00]	[0.00]	[0.00]	[0.17]
RLM ERR	597.26	315.43	503.68	274.69
	[0.00]	[0.00]	[0.00]	[0.00]

Notes to Table 1:
OLS estimation. Standard Errors presented in parenthesis and p-values in square brackets. ***, ** and * denote statistical significance at 1%, 5% and 10% respectively.

Moran's test (Moran [13]) applied to the residuals of regressions detect a positive spatial autocorrelation in the case of all models. All LM tests reject the null hypothesis of linear model but are indecisive about which spatial models optimally addresses the issue of spatial autocorrelation. Robust versions of the tests provide a clear indication of which model should be preferred in the case of the column *(d)* only. The hypothesis of spatial dependence in the dependent variable is rejected against the alternative of spatial dependence in the error.

Estimates of the spatial error model in equation (5)are reported in the first column of Table 2. Results are consistent with the evidence based on the linear model in terms of slope and magnitude of estimated coefficients. The model is used thus to investigate the issue of structural stability of coefficients. The sample is split into four regimes, using 25% quintiles of the population distribution.

Table 2: Spatial Error estimation and regime analysis

	Spatial Error Model	Population Regimes			
		(0; 1147]	***(1147;2644]***	***(2644; 5674]***	***(5674; …***
Intercept	218.450***	108.190*	142.050**	254.980***	441.860***
	30.888	62.887	57.520	59.177	42.674
Y	-54.187***	-2.820	-11.783	-4.256	-208.280***
	12.062	20.519	19.326	20.515	18.833
P	0.123***	0.366**	0.264***	0.159***	0.096***
	0.003	0.144	0.103	0.050	0.003
R	2.625***	0.242	0.658	1.562***	2.330***
	0.517	0.726	0.564	0.503	0.591
T	-255.650***	-322.870	-359.810	-691.610***	590.040***
	84.533	255.640	268.980	266.390	86.955
λ	0.688	0.615***			
	0.037	0.043			
LR	207.75	132.85			
	0.00	0.00			
AIC	19800.00	19339.00			
SpChow		491.73			
		0.00			

Notes to Table 2.

ML estimation. Standard Errors are presented in parenthesis and p-values in square brackets. ***, ** and * denote statistical significance at 1%, 5% and 10% respectively. . The value of the LR test refers to the spatial parameter of the model. The test for structural instability (SpChow) is performed according to Anselin [3].

The estimated values of the intercept monotonically increase moving from the lower to the higher regime, consistently with expectations. The slope of the income coefficient is negative in all regimes, although statistically significant only in the regime with the highest population, further supporting the explanation for the negative income effect based on house prices in large cities and inter-city commuting from suburbs. In fact, the effect is relevant in large urban agglomerations only, where house prices are, on average, substantially higher and connection with other cities outside the metropolitan area is easier than in small cities.

The effect of the population is positive and largely significant in all regimes. Furthermore, evidence indicates that the incidence on urbanized area of an increase in population is higher in less populated cities compared to more populated ones, as the coefficient estimate decreases monotonically from low to high regimes.

The evidence related to the positive effect of agricultural rent weakens when regime-specific coefficients are estimated. In particular, for both regime typologies, it is found that such a positive relationship is shown in medium/large and large urban agglomerations. Accordingly, in the presence of substantially relevant urban pressures, agricultural rents tend to follow urban size more than the opposite, as agricultural land values discount the higher probabilities of land use change in the area.

Finally, results about the transportation costs variable provide mixed indications of the overall effect. While an increase in transport costs is predicted to increase the extent of urbanization in large urban agglomerations, an opposite effect is evidenced in medium/large cities and no effect at all in small cities.

5. Conclusion

Economic theory predicts that the equilibrium city size is related to population and income positively and agricultural rents and transport costs negatively. This simple theoretical framework was initially proposed to explain variation in the spatial scale of cities in the 80s and more recent empirical evidence based on US urban agglomerations suggests that it also survived the test of time. As modern urban expansion is also characterized by sprawl and excessive soil consumption, the empirical model has been also used to disentangle the extent to which unregulated soil consumption determines the spatial scale of cities more than the interaction of market forces predicted by the model.

In this paper, this empirical framework is subject to the test of small cities and is used to explain variation in city size across municipalities in the Lombardy region, Italy. Evidence suggests that even in small cities urbanization obeys to economic and market forces, but to a limited extent only. Concerning income it is found that the relationship with size is negative and significant in larger cities only; the result is not new in the literature and is associated with the competition between neighboring cities in attracting workers. House prices are particularly high in large urban agglomerations and individuals may prefer to live in neighboring cities and to commute between cities rather than to live at the edge of the city and commute from the edge to the center.

Demographic trends are the most important determinant of city size, results indicate. However, a rise in population brings a substantially greater effect in small cities than in large ones framing the possibility that urban sprawl phenomena operate in these cities. From the land take viewpoint, the spatial concentration of economic activities in some agglomerated areas threatens the equilibrium of agricultural systems less than spreading of these activities across a network of small cities. The scope of this result and the consequent implications for urban policy is limited to the Lombardy region and cannot be generalized to the European case. Nonetheless the characters of urbanization in this region, especially regarding external pressures on the agricultural sector, are not very different from those of many other capital regions in Europe and, in general, of regions hosting large urban agglomerations such as Milan.

Among other explanations for city size variation, agricultural land rent does not appear as relevant. It is found, on the opposite, that land rents are in a positive relationship with city size. The use of current agricultural price as agricultural rents proxy explains this evidence to a large extent. Prices, in fact, also discount the value of land use change and the probability that this change happens in the next future, not only the rent from agricultural activity. Even so, the measure is the best proxy available at the municipality level, and results are robust to the exclusion of this variable from the econometric model.

Finally, city size is explained by transportation costs. Reducing the commuting cost between the edge and the center of the city increases its size. Much depend, however, on how transport costs are measured. The proxy used in this paper relates costs to the use of cars and hence this measure is artificially higher in large urban agglomerations, where public transportation weights more. This explains why the evidence in this paper indicates that size grows with increasing transport costs in larger cities.

References

[1] Anselin L, "Spatial Econometrics: Methods and Models", Kluwer Academic Publisher, 1988.

[2] Anselin, L, "Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity", Geographical Analysis, 20, 1988, 1–17.

[3] Anselin, L, "Spatial Dependence and Spatial Structural Instability in Applied Regression Analysis", Journal of Regional Science, 30, 1990, 185–207.

[4] Brueckner JK.,"Urban Sprawl: Diagnosis and Remedies", International Regional Science Review, 23, 2000, 160–71.

[5] Brueckner JK. and Fansler DA., "The Economics of Urban Sprawl: Theory and Evidence on the Spatial Sizes of Cities", The Review of Economics and Statistics, 65, 1983, 479–82.

[6] Brueckner JK. and Sridhar KS, "Measuring Welfare Gains from Relaxation of Land-Use Restrictions: The Case of India's Building-Height Limits", Regional Science and Urban Economics, 42, 2012, 1061–67.

[7] Cervero R, and Wu KL, "Sub-Centring and Commuting: Evidence from the San Francisco Bay Area, 1980-90", Urban Studies, 35, 1998, 1059–76.

[8] Deng X, Huang J, Rozelle S and Uchida E, "Growth, Population and Industrialization, and Urban Land Expansion of China", Journal of Urban Economics, 63, 2008, 96 – 115.

[9] Glaeser EL and Kahn ME, "Sprawl and Urban Growth", in Handbook of Regional and Urban Economics, Henderson VJ and Thisse JF (eds), Elsevier, 2004.

[10] LeSage J and Pace K, "Introduction to Spatial Econometrics", CRC Press, 2009.

[11] McGrath DT, "More Evidence on the Spatial Scale of Cities", Journal of Urban Economics, 58, 2005, 1–10.

[12] McMillen DP, "Testing for Monocentricity", in A Companion to Urban Economics, Arnott RJ and McMillen DP, Blackwell Publishing Ltd, 2007.

[13] Moran PAP, "Notes on Continuous Stochastic Phenomena", Biometrika, 37, 1950, 17–23.

[14] Paulsen K, "Yet Even More Evidence on the Spatial Size of Cities: Urban Spatial Expansion in the US, 1980–2000", Regional Science and Urban Economics, 42, 2012, 561–68.

[15] Song W, Baiming C and Ying Z "Land-Use Change and Socio-Economic Driving Forces of Rural Settlement in China from 1996 to 2005" Chinese Geographical Science, 2013, 1–14.

[16] Spivey, C "The Mills—Muth Model of Urban Spatial Structure: Surviving the Test of Time?", Urban Studies, 45, 2008, 295–312.

[17] Wassmer, RW "The Influence of Local Urban Containment Policies and Statewide Growth Management on the Size of United States Urban Areas", Journal of Regional Science, 46, 2006, 25–65.

[18] Wheaton, WC "A Comparative Static Analysis of Urban Spatial Structure", Journal of Economic Theory, 9, 1974, 223–37.