Income, distance and spatial effects in the Monocentric model. The Barcelona Metropolitan Area case.

Rafa Madariaga
Joan Carles Martori
Ramon Oller
Department of Economics and Business
Universitat de Vic (Barcelona)

Abstract

This paper focuses on the income distribution of households in Barcelona Metropolitan Area. For this purpose we use the Monocentric model. As the basic model does not have direct implications for this distribution, we survey the extensions of the model that have been used in empirical literature. One of the most promising ways is to introduce externalities in the decision process; they can result directly from exogenous amenities (natural traits of urban area) or be created directly by other agents' decisions. In this case, a spatial lag model and spatial error model are suited for the empirical purposes. We present evidence that any model with spatial effects improves significantly the econometrical results.

Keywords: Monocentric model, income-distance relationship, spatial effects.

1. Introduction

The Monocentric city model is the cornerstone of urban economics since its formulation in the decade of 1960 by Alonso, Muth and Mills. It is the first theoretical model in urban economics and it has generated a huge quantity of work in theoretical and empirical grounds. The implications of monocentric model, especially for the relations between distance to the Central Business District (CBD) and population density, housing prices, land rent and capital/land ratio are widely known and have been tested many times for a great number of cities and countries. Last references in this area are Baum-Snow (2007) who presents a version with radial commuting highways to analyze new forms of commuting. Spivey (2008) examines the viability of some basic predictions of the model of city structure for modern cities. McDonald (2009) introduces a new perspective of traffic congestion in the monocentric model. The distribution of population in an urban area and its evolution over time has been extensively studied. But, what are the patterns of income distribution in urban areas? How they evolved over time? These are not just theoretical questions, but interesting questions for policy making.

Less much research has been carried out on the relation between distance and income. In a recent review of empirical evidence on the Monocentric model predictions by McMillen (2010) nothing is said about the distance-income relation. As we will show, the most basic version of the model does not have clear implications for the relation between household income and distance. For the model to have predictions about this relation is necessary to add more assumptions.

Glaeser et al. (2008), show different patterns for “old” and “new” cities in USA. Following LeRoy and Sonstelie (1983) they develop a model with different transportation modes and analyse the effects of them on the spatial income distribution of households. Bartolome and Ross (2003, 2007) following Tiebout (1956) classical model have investigated the importance of jurisdictions having different public services and tax levels in the income-distance relation. They show that the model is capable to generate “income mixing” equilibriums. Other literature strand (Mieszkowski et al. 1993, Mills et al. 1997, Brueckner et al. 1999) argues that externalities linked with neighbourhood characteristics (Hills, landscapes, coastlines, crime, racial composition, pollution, and so on) can alter substantially the relation between income and distance.

The Monocentric model is subject to easy criticism because is static and because one of its main assumptions is that there is a CBD to where all households have to commute for working. Although the Monocentric model has been challenged by more realistic assumptions that have leaded to the formulation of new polycentric models it is still widely used for empirical purposes. And as Mills (2000) puts it, “the chimp still types”. Our opinion is that the broad use of Monocentric model is explained by two main reasons. First, its mathematical and statistical structure is simple enough to be very flexible and can be tested in different ways with easily available data in developed countries. Second, the generation of subcentres and the equilibrium for polycentrical structures requires more mathematical and statistical sophistication in the models that introduces more difficulties for the empirical work. Following Mori (2006), to endogenize the formation of centres and sub-centres requires at least one of three elements: Space heterogeneity, non-market externalities or imperfect competition.

Considering that the distance-income relation has been less studied than other implications of Monocentric model and that there are no works for Spain, Catalonia, or Barcelona Metropolitan Area (BMA), we focus only on this model. The main objective of this paper is to analyse and describe the income-distance profile in the BMA. Our research is twofold. In theoretical grounds we present a general framework and the extensions of the Monocentric model that have been developed to analyse distance-income profiles. Empirically, first we estimate the best-fit statistical function between income and distance. Afterwards, we specify and test an econometric model that takes account of the externalities that are present in the distribution of income. We consider two kinds of models: spatial lag model and spatial error model. Although we do not carry out an analysis of the origin of these externalities (further research is needed), we present robust evidence of their importance to understand the urban income structure.

The paper is organised in the following manner: the next section overviews the basic Monocentric model, focusing on the relation between income and distance. We also consider the extensions that the literature have put forward to improve the model to explain empirical patterns. In the third section we explain our empirical strategy and the main results we have obtained testing for multiple functional forms. The fourth section contains our conclusions and future research questions.

2. Theoretical overview

In this section we set out the basic model that relates distance with income. We consider a circular city of radius x_m which contains a CBD in the centre. The city population N is assumed exogenous. This implies that we take into account only one city and that we are avoiding the interaction between the cities. This corresponds to the close-city model. Alternatively, the population could be determined endogenously. This corresponds to the open-city model, where individuals are indifferent between the city and any other place. In the first case, the utility level of the city is endogenous and differences between city levels will generate migrations; in the second, the utility level is exogenously fixed by the city system and migrations equalize the utility level.

All job opportunities are located in the CBD to where all households must commute. Households have an exogenous income y, obtained working in CBD. Households preferences are expressed by an utility function u=U(c,L) which depends on the consumption of a composite good, c, and on the lot size of housing, denoted by L. Note that this assumption implies that what households value is the interior space of the housing. All the attributes of the housing are considered jointly with the composite good. The usual assumptions on the utility function are applied. Composite good is used as numeraire so its price is one. We specify a quasilinear or consumer surplus utility function:

(1) $U = U(c,L) = c + V(L)$

Each household with income y_i have to pay the transportation costs for commuting depending on the distance T(x), and the cost of housing. This cost is the product of the lot size L by the price or bid rent of a unit of housing. The rent-bid function for unit of land for a household depends on the distance and on the income, R(x,y). With these assumptions, the consumption of the composite good is:

(2) $c = y – T(x) – L\, R(x,y)$

So, in equilibrium, the utility level of the household is:

(3) $U(x,y) = y – T(x) – L\, R(x,y) + V(L)$

Each household chooses a location which maximizes utility subject to the budget constraint. For the household to be indifferent between two locations, the utility level must be the same (u_0). We can calculate the individual bid-rent function for a household:

(4) $R(x,y) = [y – T(x) + V(L) – u_0]/L$

The interaction of individual bid-rent functions in the housing market sorts households at different distances of the city centre. By the envelope theorem it can be demonstrated that in every location, the household with the greater value of the bid-rent function will outbid the other households. The equilibrium location is where the individual bid-rent function is tangent to the market bid-rent function. This equation implies that:

(5) $\partial R(x,y)/ \partial x = - T'(x)/L$

This expression is the workhorse of urban equilibrium and is found, with alternative formulations, in the most of the works on density gradients, land rent gradients and floor/area gradients[1]. In this form, it is not related with households income levels; more assumptions are needed to obtain any prediction.

For this purpose, it is necessary to relate the equation with the income level for each household. There are at least three ways that have been used in the literature. The first one is to consider that both marginal cost of commuting and lot size hinge on income. Furthermore, consider that marginal cost of commuting is a constant t, related to income. Under such assumptions:

(6) $\partial R(x,y)/ \partial x = - t(y)/L(y)$

This equation makes clear two opposite forces in location by income. The high housing consumption of high-income households makes them more attracted by farther locations where housing is cheaper. However, high-income households also have a high opportunity cost of time and then high marginal commuting cost. In other words, the slope of rent-bid functions as income increases depends on the ratio t/L. If it is decreasing with income, rich households tend to live in the suburbs. If it increases, the poor households tend to live in the suburbs. Formally, the derivative of this slope with respect to income and some algebra yields:

(7) $\partial^2 R(x,y)/ \partial x \partial y = -t/yL(y)\ [\varepsilon_y^{\,t} - \varepsilon_y^{\,L}]$

Where ε_y^t and ε_y^L are the income elasticities of marginal commuting costs and lot size, respectively. If $\varepsilon_y^t > \varepsilon_y^L$ the slope of rent-bid functions decreases when income increases: low-income households live near the centre and high income households in the suburbs. In the opposite case, the location pattern reverses. This result assumes that marginal commuting cost is different for different income households. If we suppose only one mode of transportation and that the main cost of commuting is time, equation (7) is:

(8) $\partial^2 R(x,y)/ \partial x \partial y = -[t/L]\ [1- \varepsilon_y^{\,L}]$

In this case, low-income households will live in centre if the income elasticity of housing is greater than one. In this case, the rich households tend to live in suburbs.

The third way is to consider the use of disposable time. Each household has a unit of time and uses it for earning income which, after paying transport cost, allows him to pay the bid rent for the housing lot. The transportation costs can be expressed as a function of income. Income is the opportunity cost of commuting. So, transportation cost is the product of unit transportation cost t (time/km), a constant, by distance to the CBD, x (Km) by income y (€/time):

(9) $T(x) = tyx$

In this case, equation (5) implies that the rent-bid function has a slope:

---

[1] To close the model, two additional conditions are required: the bid rent at the edge of the city must be equal to the agricultural rent, $R(x_m,y)= r_a$, and the whole population N, have to be located in the city: $\pi x_m^2 = NL$.

(10) $\partial R(x,y)/ \partial x = -ty/L$

The individual bid-rent function is decreasing function with the absolute value of slope depending directly on unit transportation cost and income level and inversely on lot size. Bartolome and Ross (2007) consider equal housing lots; in this case, the bid-rent functions of those with higher income will be steeper and they will outbid poor households from central locations. This case generates income-decreasing gradients which parallel the gradients for density, housing prices and capital/land ratios.

Glaeser *et al.* (2008), present strong evidence of the centralization of poor households in American cities. Furthermore, they estimate income elasticities of the transportation costs and lot size. With these estimations the model does not explain the observed pattern. In their view, the most accurate way to explain it in the Monocentric model is to introduce different transportation modes. Following LeRoy and Sonstelie (1983) they develop a model with different transportation modes with different transportation costs. They argue that if there is one transportation mode, the ratio t/L rises with income so the high-income households would tend to live in the centre. However, if rich households can switch to a faster transportation mode (say car), the marginal commuting cost can decrease for high-income households but remain for low-income households that cannot afford to pay for the more expensive transportation mode. This implies a reversal of the location pattern. The model may generate different equilibriums, depending on the availability (cost) of the different modes.

Bartolome and Ross (2003, 2007) develop a model with different jurisdictions that have different public services and taxes. They broaden the utility function with the public service supplied by each jurisdiction as a scalar. They show that this model can generate "income mixing" equilibrium opposite to "income sorting" equilibriums. At the jurisdiction border the bid-rent function presents a discontinuity. Summing up, the different fiscal local regimes may attract rich households to the jurisdictions that surround the central city.

Brueckner *et al.* (1999) present an amenity-based model of location by income. This model formalizes previous claims on the relevance of urban space characteristics. They consider exogenous and endogenous amenities. The first category consists on amenities that depend on historical (monuments, buildings, palaces, bridges) and physical (rivers, hills, coastlines) urban traits. Households can value the proximity to these exogenous amenities so this can attract households. The second category consists on endogenous amenities related with neighbourhood characteristics. Between them, they quote restaurants, theatres, and sport facilities. In either case, the distribution of amenities affects the distribution of households and then, the income-distance profiles. The model can be extended to bear also with disamenities; space characteristics that affect negatively the utility of households. They can be also exogenous or endogenous. In either case, the model should be enlarged to include the amenities in the utility function. Let's assume exogenous amenities that could be represented by a scalar depending on location, A(x).[2]

(11) $U(x,y) = y - T(x) - L\,R(x,y) + V(L) + A(x)$

The household maximizes utility subject to the budget constraint and obtains a utility level u_0. Indifference between locations implies a bid-rent function:

[2] It can be thought as an index number with different weighted attributes.

$$(12) \quad R(x,y) = R(x,y) = [y - T(x) + V(L) + A(x) - u_0]/L$$

The slope of this function is:

$$(13) \quad \partial R(x,y)/ \partial x = - T'(x)/L + A'(x)/L$$

The second term of this equation is the effect of amenities on the slope of bid-rent function. The sign of this term depends on the distribution of amenities in the city. For example, If amenities were concentrated in the centre, A'(x) will be negative and great. Where the model predicts the tendency of rich households to locate near the centre, this location pattern of amenities will reinforce the tendency[3]. If the amenitites were concentrated far from the centre, they will create an additional force that will reinforce the rich households to decentralize. Depending on the location patterns discussed above, it can reinforce or counteract the locational forces.

The amenities can be endogenized making them depending on income level. Glaeser (2008) sets out an amenity function related to the average income surrounding each location, $A(\hat{y}(x))$. The average income can be calculated using a variable number of space units close to each one. In this model, the condition (13) can be rewritten:

$$(14) \quad \partial R(x,y)/ \partial x = - T'(x)/L + A'(\hat{y}(x))\ \hat{y}'(x)/L$$

The second term at the right side of the equation is the effect of amenities, but in this case, it depends also on the decision undertaken by other households. Notice the sign of the second term depends also on the spatial distribution of average income $\hat{y}'(x)$. In the empirical work, it creates an externality: the individual decision depends on prices, distances and housing lots but is affected also by the decision of the other households. This is the main hypothesis we test in this paper.

3. Econometric approach

Our empirical strategy goes as follows: first we look for the best model for the relation between income and distance. We use disposable research of the relation between distance and density as a guide. To avoid miss-specification problems we have conducted also a non-parametrical estimation by means of Gaussian Kernel function. The results show a very weak relation. Then we introduce spatial effects to take account of the externalities that can be present in the household location decision, following the theoretical approach that has been reviewed in the previous section. Two models with spatial effects are tested: spatial error model and spatial lag model.

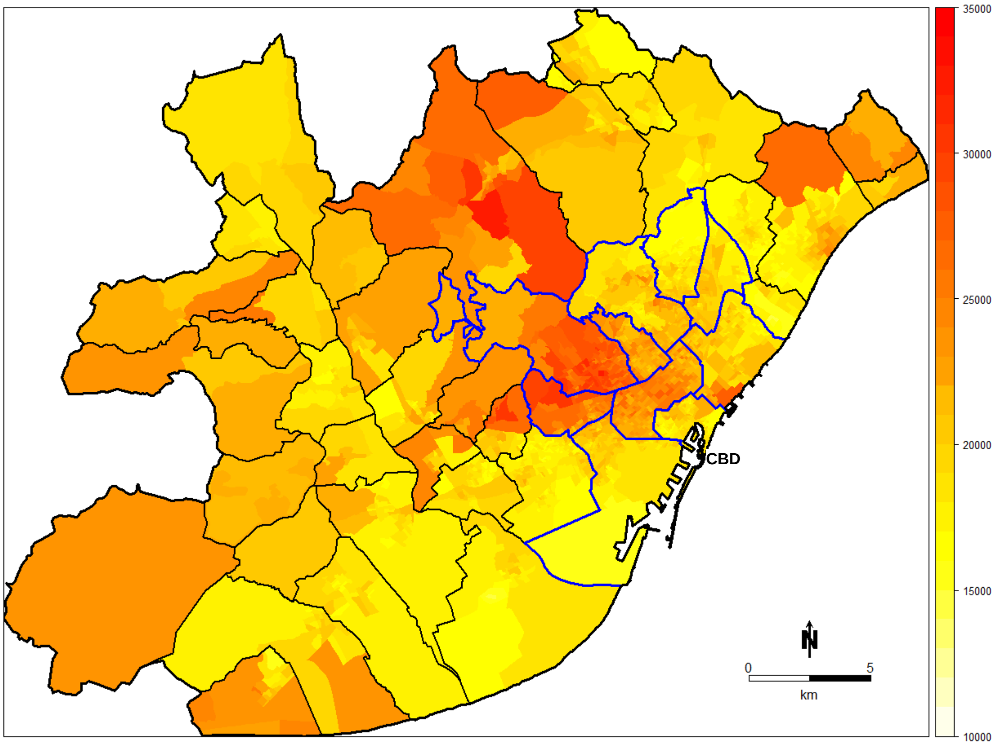
The data set for the empirical analysis is an estimation of the average wage income for 2001, for each one of the 2500 census tracts in which the 35 municipalities of the BMA is divided. The data have been obtained matching data from Census and the Wage Structure Survey

[3] This is the hypothesis that Brueckner et al. (1999) argue for the different patterns in Paris and Detroit.

(WSS) carried out by the INE (National Statistical Institute)[4]. From the census we have the total number of wage earners for 72 categories of occupation and economic sector. We have assigned for each category the Annual Average Wage estimated by the WSS. This procedure yields an estimation of the Total wage income for the corresponding census tract. Weighting this Total wage income by the number of wage earners for whom we have wage income information yields an estimation of the average wage income by census tract.

Figure 1 presents the average wage income for the BMA. It includes the boundaries of each one of the 35 municipalities we analyze. The boundaries of the 10 districts in which the city of Barcelona is divided are also plotted. Following previous research by Martori and Suriñach (2002), the Central Business District (CBD) is located in a census tract including the Barcelona Harbour.

Figure 1: Average wage income by census tract. BMA. 2001



A first look at Figure 1 does not suggest a clear relation between distance and income. Moreover, the spatial distribution of income suggests spatial dependence. That is, spatial clusters with high income values and clusters with low income can be observed. Thus it seems meaningful to test first a model which relates income and distance, and then to extend the empirical model with some spatial structure.

Monocentric urban density analysis has received considerable attention from two disciplines: urban geography and regional science, where it has had both theoretical and empirical applications. The classical study undertaken by Colin Clark (1951) has generated an extensive body of literature dealing with empirical implementations for a wide range of metropolitan areas and cities, in different countries and at different times. Here, we analyse seven functional forms that originate from both theoretical and empirical models. Some of these

[4] For further details on data set construction, and the Spanish definition of census tract, see Madariaga et al. 2009.

functions have been used in traffic planning studies, for example Tanner (1961) and Smeed (1963), while others have been employed in theoretical models of the housing market in the monocentric case (Muth 1969, 1971). The generalisation of the functional form and the comparison of results are the work of Casetti (1973), McDonald and Bowman (1976), Kau and Lee (1976a, 1976b), Zielinski (1979), Anselin and Can (1986), Smith (1997), Wang and Zhou (1999), Bunting et al. (2002, 2004), and Filion et al. (2004, 2010). The functional form of urban population density is not unique and this implies that a selection process must be adopted in each case. Martori et al. (2002) have contributed to this strand presenting a general framework of the functional forms that have been used. In Table 1 we present a summary of the classical and most frequently used functional forms.

We start with the Zielinsky model. It has been showed that many other functional forms are nested to this general formulation. The relation between income and distance is supposed to be:

(15) $$y_i = \alpha x_i \exp\left(\beta_2 x_i + \beta_3 x_i^2\right)$$

Where y_i is income and x_i is the distance to the CBD. Taking logarithms, the model can be estimated as follows:

(16) $$ln y_i = \beta_1 + \beta_2 ln x_i + \beta_3 x_i + \beta_4 x_i^2$$

The OLS estimation of this model allows us to choose the better specification between the set of nested models. We based our selection on the F statistic and the log-ratio test statistic. The results are shown in Table 2. It can be seen that using any criterion, the Zielinski model yields the best estimation. The results of the ordinary least squares (OLS) estimation process are given in Table 1.

Table 2: Selection model results

Model	**AIC**	**LR test**	**Adjusted R^2**	**F test**
Zielinski 1	-2567.3	-	0.17	-
Newling 1 (β_1=0)	-2350.6	218.6***	0.10	228.0***
Zielinski 2 (β_2=0)	-2270.7	298.6***	0.07	316.4***
Aynvarg (β_3=0)	-2262.5	306.8***	0.07	325.7***
Clark (β_1= β_3=0)	-2251.6	319.7***	0.06	170.1***
Tanner (β_1=β_2=0)	-2185.0	386.3***	0.04	208.3***
Smeed (β_2= β_3=0)	-2264.3	307.0***	0.07	162.9***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 ' ' 1

As is usual in this topic, we have considered that inferences can be affected by the presence of heteroskedasticity. The result for the Breush-Pagan test in Table 1 confirms this fact. Since the sample size is large, the consistency property assures that it does not affect the proper estimation of the parameters. However, heteroskedasticity has consequences on the inferences of the parameters unless we use a consistent estimate of the covariance matrix of the coefficients. We studied the significance of the coefficients using several consistent estimates of the covariance matrix proposed by White (White 1980) and MacKinnon and White (1985). All the coefficients (not reported) result to be highly significant.

TABLE 1. POTENTIAL INCOME - DISTANCE FUNCTIONAL FORMS

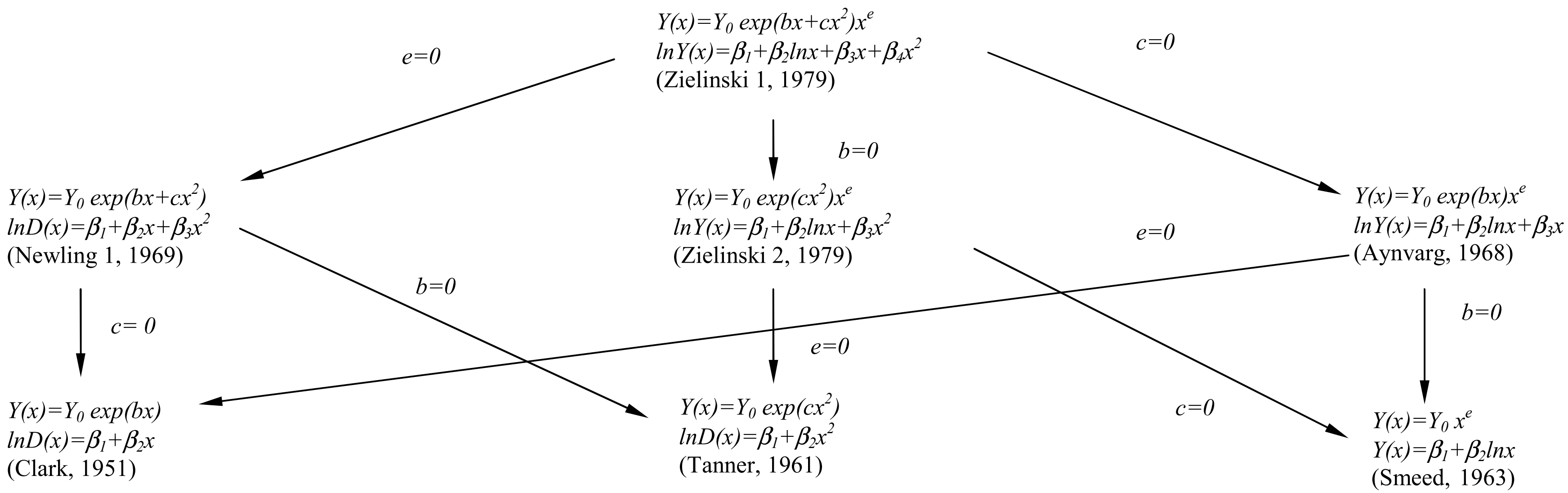


Table 1: Estimation results of the Zielinski's model

	OLS	LAG
β_0	9.73134***	1.01975***
β_1	0.77559***	0.10673***
β_2	-0.23041***	-0.03159***
β_3	0.00636***	0.00087***
Rho		0.89437***
BP	96.27***	76.24***
LM_{lag}	4175.15***	
LM_{error}	4153.21***	
RLM_{lag}	30.64***	
RLM_{error}	8.70 **	
R^2	0.17	0.92
AIC	-2567.3	-6132.6

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 ' ' 1

The model selection procedure followed above requires that one of the models has to be the best-fit functional form. Obviously, this procedure has the power to choose between some specifications but it does not consider all of them. A miss-specified model could result in completely wrong inferences. Alternatively, a nonparametric model introduces more flexibility in the relationship between the variables and does not preclude a closed structure in advance. This issue has received increasing attention in econometrical literature applied to urban economics, see for instance McMillen (2010) and McMillen and Redfearn (2010).

For the sake of completeness of the analysis of the relationship between income and distance to CBD, we have developed a kernel regression. Our goal is to compare the nonparametric model and the Zielinski's model, and assure that the parametric model is well specified. We consider the next nonparametric model:

(17)

$$lny_i = f(x_i) + u_i$$

The f function is completely unspecified and u is an usual error term. The model has been estimated considering a Gaussian kernel function,

(18)

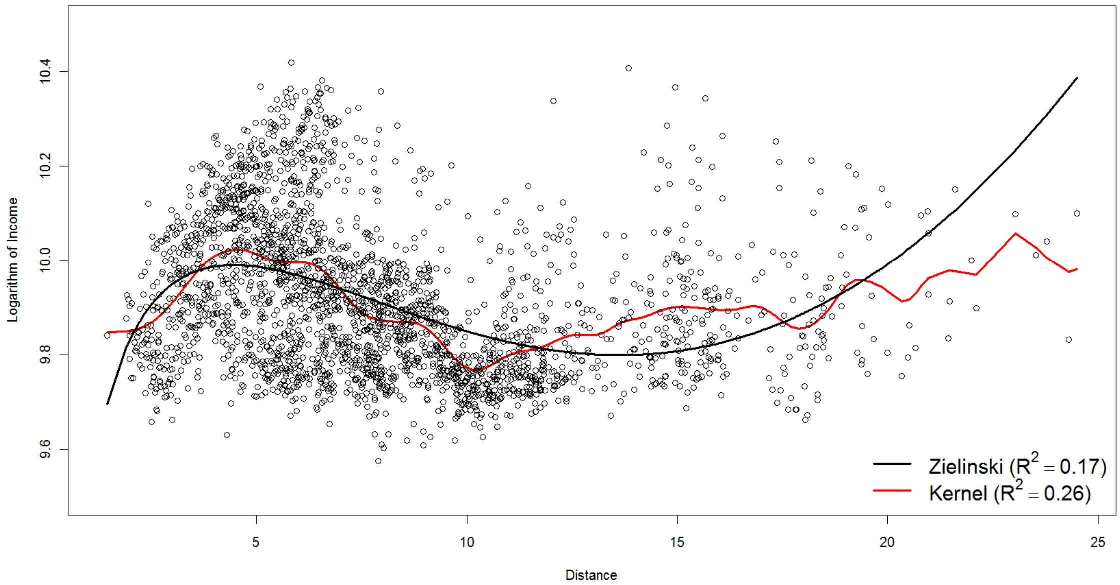
$$ln\hat{y}(x) = \frac{\sum_{i=1}^{n} \phi\left(\frac{x_i - x}{h}\right) lny_i}{\sum_{i=1}^{n} \phi\left(\frac{x_i - x}{h}\right)}$$

where ϕ represents the standard normal density function and h is the bandwidth or window size which controls the degree of smoothing of the estimated curve. To pick the window size we have considered a Monte Carlo cross-validation method (Picard et al. 1984, Arlot et al. 2010) with a training fraction of 90% of the data and 10 training sets. After comparing the average of the cross-validated mean-square errors (MSE) of all the different considered bandwidths, the optimal bandwidth was obtained at the value *h=0.4*.

In Figure 2 the parametric and the nonparametric estimations are plotted. The models imply three different gradients for three different zones. Around the CBD, for distances smaller than 5 Km, the relation between distance and income is positive and steep. It points that high income households tend to locate further the CBD, but inside the Barcelona municipality. Between 5 and 12 Km, the relation turns negative. In this second crown the lower income households are located, and more distance to the CBD implies lower wage income. The figure shows also that the poorest census tracts are located between 7 and 12 Km far from the CBD. Further than 12 Km the relation turns positive another time. But the gradient grows slower than in the first zone. Although both curves are very similar, the parametric model indicates steeper gradients in census tracts at distances above 20 kilometres. One important conclusion can be obtained. The relation between income and distance is not monotonic.

As a measure of goodness of fit, we also give in Figure 2 the square of the Pearson's correlation coefficient (R^2) between the log of the income and the fitted values obtained from each model. We observe that the nonparametric model fits better than the parametric one. This is a typical behaviour in such comparisons, even if the parametric model is correct.

Figure 2: Income-distance estimated values.



To further test the parametric and the nonparametric models, we have used a bootstrap approach. If the Zielinski's model is right, the differences between the Mean Square Errors in both curves, $U=MSE_{Zielinski}-MSE_{Kernel}$, should be approximately zero. Since we do not know the distribution of U under the null hypothesis, we have generated resampled data sets using the parametric model. In each step we have simulated new responses from the parametric fitted model given in Table 1 and then we reestimated the parametric and the nonparametric models (using cross-validation to pick an optimal bandwidth) and computed the statistic U. We repeated this process 400 times. In 399 replications the resampled value of U was greater than the observed value of U=0.0021.

The conclusion of this analysis is that the Zielinski's model has to be considered appropriate to describe the "simple" relation between income and distance. But as results in table 2 shows, it explains a small fraction of the relation.

Different reasons conduct us to introduce spatial effects. First, the empirical results that have been presented imply that the relation between distance and income is weak, and that the distance to the city centre explains a small proportion of income variability. Even more, when we use non-parametrical estimation, a misspecification problem appears clearly. Second, the theoretical analysis presents two main reasons to introduce spatial effects. If there is spatial interaction it is not possible to hold the hypothesis of independence between the observed variables. It seems reasonable to think that when families undertake their location decision, they take account of the level of income, i.e. the decision of other families. In the same way, there are factors located in "other spaces" that may have important influence in the location decision. Both arguments point to the inclusion of what has been called interacting agents or social interaction (Anselin, 2002). These arguments have been considered into the econometrical specification in two ways. The first one is referred to as spillover model or spatial lag model; it takes account that when a household with an income y_i decides the location it considers the location of other households with different values of income. The location of other households affects its objective function. An alternative way for considering this specification is the presence of amenities, neighbourhood characteristics. In the spatial error model, the agent's decision variable is not directly affected by the location chosen by other agents, but only indirectly. There is something, scarcely specified, that affect the decision of all the households.

Mur and Angulo (2009) give a recent and detailed discussion of different strategies to detect the suitable form of spatial autocorrelation. In the present study we have considered the *Specific-to-General* strategy. We computed the Lagrange Multiplier tests for spatially lagged dependent variable (LM_{lag}) and for error dependence (LM_{error}) and their robust versions (RLM_{lag} and RLM_{error} respectively). Table 1 shows that all of them are significant and indicate that the inclusion of spatial structure in the model is needed. The decision rule in this situation is to choose the spatial model with the most significant Lagrange Multiplier test. The result is only clear for the robust version; the RLM_{lag} test is highly more significant than the RLM_{error} test. Thus we estimate a spatial autoregressive model.

Table 1 shows the results of the likelihood estimation of the spatial autoregressive model. The coefficients are highly significant and the new model clearly improves the Zielinski's model in terms of the AIC statistic. As a measure of goodness of fit we considered R^2 as the square of the Pearson's correlation coefficient between the log of the income and the fitted values. We observe in Table 1 that the estimation improves substantially in relation with the Zielinski's model. The spatial Breush-Pagan test confirms again the presence of heteroskedasticity.

4. Conclusions

Monocentric model has been extensively used to analyse the distribution of population in an urban area. Density gradients are the main and most famous result. Housing prices, land rent and capital/land ratio also have been studied with this set-up. The household income distribution has raised less interest; less work has been done in this field. The basic model needs to be enlarged to give predictions about this distribution. In this paper we have presented the extensions of the model that have been developed to cover this issue. One of the strands of the literature has focused on the (exogenous or endogenous) amenities that may have influence on the location decisions by households. We have focused on the average level of wage income around a census tract as it can create a kind of externality that affects the objective function.

The empirical analysis of the Zielinski model between income and distance shows a significant but weak relationship. According to the theoretical framework, we expect that income in a census tract depends on distance but also on the income in the surrounding area. Thus, we expect an improvement of the results when we introduce some spatial structure to the model. In this sense, when we introduce a spatial lag structure to the Zielinski model, the R^2 and the AIC statistics increase substantially. Other spatial structures are also possible. Our aim in future work is to extend the econometric analysis to the presence of the both lag and error effects with the SARAR model. We also plan to introduce the effect of spatial heterogeneity by means of a Weighed Geographical Regression or other similar models. We would like to investigate deeper the heteroscedasticity problem with heteroscedasticity and autocorrelation covariance (HAC) matrix estimation.

References

ANSELIN, L. Under the Hood Issues in the Specification and Interpretation of Spatial Regression Models. *Agricultural Economics*, 2002, vol. 27, no. 3, pp. 247-267.

ANSELIN, L.; and CAN, A. Model Comparison and Model Validation Issues in Empirical Work on Urban Density Functions. *Geographical Analysis*, 1986, vol. 18, no. 3, pp. 179.

ARLOT, S.; and CELISSE, A. A Survey of Cross-Validation Procedures for Model Selection. *Statistics Surveys*, 2010, vol. 4, pp. 40-79.

BAUM-SNOW, N. Suburbanization and Transportation in the Monocentric Model. *Journal of Urban Economics*, 2007, vol. 62, no. 3, pp. 405. ISSN 0094-1190.

BRUECKNER, J. K.; THISSE, J. F.and ZENOU, Y. Why is Central Paris Rich and Downtown Detroit Poor? An Amenity-Based Theory. *European Economic Review*, 1999, vol. 43, no. 1, pp. 91-107.

BUNTING, T.; FILION, P.and PRISTON, H. Density Gradients in Canadian Metropolitan Regions, 1971-96: Differential Patterns of Central Area and Suburban Growth and Change. *Urban Studies*, 2002, vol. 39, no. 13, pp. 2531.

BUNTING, T. E. Decentralization Or Recentralization? A Question of Household Versus Population Enumeration, Canadian Metropolitan Areas 1971-1996. *Environment and Planning A*, 2004, vol. 36, no. 1, pp. 127-148.

CASETTI, E. Testing for Spatial-Temporal Trends: An Application to Urban Population Density Trends using the Expansion Method. *Canadian Geographer/Le Géographe Canadien*, 1973, vol. 17, no. 2, pp. 127-137.

CLARK, C. Urban Population Densities. *Journal of the Royal Statistical Society.Series A (General)*, 1951, pp. 490-496.

DE BARTOLOME, C. A. M.; and ROSS, S. L. Community Income Distributions in a Metropolitan Area. *Journal of Urban Economics*, 2007, vol. 61, no. 3, pp. 496-518.

DE BARTOLOME, C. A. M.; and ROSS, S. L. Equilibria with Local Governments and Commuting: Income Sorting Vs Income Mixing. *Journal of Urban Economics*, 2003, vol. 54, no. 1, pp. 1-20.

FILION, P., et al. Canada-US Metropolitan Density Patterns: Zonal Convergence and Divergence 1. *Urban Geography*, 2004, vol. 25, no. 1, pp. 42-65.

FILION, P., et al. Intensification and Sprawl: Residential Density Trajectories in Canada's Largest Metropolitan Regions. *Urban Geography*, 2010, vol. 31, no. 4, pp. 541-569.

GLAESER, E. L. *Cities, Agglomeration, and Spatial Equilibrium.* Oxford University Press, 2008.

GLAESER, E. L.; KAHN, M. E.and RAPPAPORT, J. Why do the Poor Live in Cities the Role of Public Transportation. *Journal of Urban Economics*, 2008, vol. 63, no. 1, pp. 1-24.

KAU, J. B.; and LEE, C. F. The Functional Form in Estimating the Density Gradient: An Empirical Investigation. *Journal of the American Statistical Association*, 1976a, pp. 326-327.

KAU, J. B.; and LEE, C. F. Functional Form, Density Gradient and Price Elasticity of Demand for Housing. *Urban Studies*, 1976b, vol. 13, no. 2, pp. 193-200.

LEROY, S. F.; and SONSTELIE, J. Paradise Lost and Regained: Transportation Innovation, Income and Residential Location. *Journal of Urban Economics*, 1983, vol. 13, no. 1, pp. 67-89.

MACKINNON, J. G.; and WHITE, H. Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties. *Journal of Econometrics*, 1985, vol. 29, no. 3, pp. 305-325.

MADARIAGA, R.; MARTORI, J.C.; OLLER, R. The spatial inequality of wage income in Barcelona Metropolitan Area. 2009, Documents de Treball, Universitat de Vic, Spain.

MARTORI, J. C.; and SURINACH, J. Classical Models of Urban Population Density. the Case of the Barcelona Metropolitan Area. *Nederlandse Geografische Studies*, 2002, no. 303, pp. 109-123.

MCDONALD, J. F. Calibration of a Monocentric City Model with Mixed Land use and Congestion. *Regional Science and Urban Economics*, 2009, vol. 39, no. 1, pp. 90-96.

MCDONALD, J. F.; and BOWMAN, H. W. Some Tests of Alternative Urban Population Density Functions. *Journal of Urban Economics*, 1976, vol. 3, no. 3, pp. 242-252.

MCMILLEN, D. P.; and REDFEARN, C. L. Estimation and Hypothesis Testing for Nonparametric Hedonic House Price Functions. *Journal of Regional Science*, 2010, vol. 50, no. 3, pp. 712-733.

MCMILLEN, D. Issues in Spatial Data Analysis. *Journal of Regional Science*, 2010, vol. 50, no. 1, pp. 119.

MIESZKOWSKI, P.; and MILLS, E. S. The Causes of Metropolitan Suburbanization. *The Journal of Economic Perspectives*, 1993, vol. 7, no. 3, pp. 135-147.

MILLS, E. S.; EPPLE, D.and OATES, W. E. A Thematic History of Urban Economic Analysis [with Comments]. *Brookings-Wharton Papers on Urban Affairs*, 2000, pp. 1-52.

MILLS, E. S.; and LUBUELE, L. S. Inner Cities. *Journal of Economic Literature*, 1997, vol. 35, no. 2, pp. 727-756.

MORI, T. Monocentric Versus Polycentric Models in Urban Economics. *KIER Working Papers*, 2006.

MUR, J.; and ANGULO, A. Model Selection Strategies in a Spatial Setting: Some Additional Results. *Regional Science and Urban Economics*, 2009, vol. 39, no. 2, pp. 200-213.

MUTH, R. F. The Derived Demand for Urban Residential Land. *Urban Studies*, 1971, vol. 8, no. 3, pp. 243-254.

MUTH, R. F. *Cities and Housing*. Univ. of Chicago Press, 1969.

PICARD, R. R.; and COOK, R. D. Cross-Validation of Regression Models. *Journal of the American Statistical Association*, 1984, vol. 79, no. 387, pp. 575-583.

SMEED, RJ. The Effect of some Kinds of Routing Systems on the Amount of Traffic in Central Areas of Towns. *Journal of the Institution of Highway Engineers*, 1963, vol. 10, no. 1, pp. 5-26.

SMITH, B. E. A Review of Monocentric Urban Density Analysis. *Journal of Planning Literature*, 1997, vol. 12, no. 2, pp. 115.

SPIVEY, Christy. The Mills-Muth Model of Urban Spatial Structure: Surviving the Test of Time?. *Urban Studies*, 2008, vol. 45, no. 2, pp. 295. ISSN 0042-0980.

TANNER, J. C.; and Road Research Laboratory (Londres). *Factors Affecting the Amount of Travel.* HM Stationery Off., 1961.

TIEBOUT, C. M. A Pure Theory of Local Expenditures. *The Journal of Political Economy*, 1956, vol. 64, no. 5, pp. 416-424.

WANG, F.; and ZHOU, Y. Modelling Urban Population Densities in Beijing 1982-90: Suburbanisation and its Causes. *Urban Studies*, 1999, vol. 36, no. 2, pp. 271.

WHITE, H. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 1980, vol. 48, no. 4, pp. 817-838.

ZIELINSKI, K. Experimental Analysis of Eleven Models of Urban Population Density. *Environment and Planning A*, 1979, vol. 11, no. 6, pp. 629-641.