Local public goods and industrial location: A case study for the electronics industry in Madrid

Abstract:

We use geo-referenced establishment data to estimate parameters of a Gibbs model. The statistical model is used to decompose the conditional intensity of the spatial point process into trend and interaction components. The trend captures covariates related to firms' costs or profits, plus distance to public transport infrastructures, to technical universities and to cultural and recreational facilities. The ability to specify a *Geyer* interaction component captures the existence of additional spillovers providing a deeper insight into inter-establishment spatial dynamics than any previously published methods.

The results challenge some of the outcomes of the inter-urban industrial location literature, confirming that spatial aggregation compromises results in studies of business location. Firms' location decisions are dominated by site costs, with transportation costs being much less of a consideration; geographical knowledge spillovers are confirmed for large establishments; and amenities are unlikely to be important location factors.

JEL classification: C49, C51, H41, H54, R38

Keywords: industrial location, Gibbs models, local public goods**.**

1. INTRODUCTION

There are many open research questions at the intersection of regional science, local public finance and urban economics. The aim of this paper is to analyze the distribution of firms in the electronics industry within contemporary metropolis, highlighting the significance of certain local public goods and raising the importance of spatial considerations in policy evaluation. The distribution of public activities can have a substantial impact on private sector locational choices, as benefits from public services may not accrue by the same amount to all residents in a region. Many public sector activities have specific locations in urban areas and it is often of value to firms or individuals to be located in close proximity to such facilities because it is necessary or desirable to take advantage of the services they provide. A sound economic model of location decisions based on geo-referenced data permits to emphasize the relationship between a firm's location and their proximity to certain local public goods. Do local public goods affect the urban structure, or more specifically the location of firms? Can firms' locations reveal any information regarding their preferences for certain local public goods such as road infrastructures, public services that facilitate employee commuting, or geographical knowledge spillovers? These are some of the questions addressed in the paper.

Spatial aspects of local public goods are seldom considered explicitly, as a consequence of the analytical complexity involved. Most of the studies in which the implications of spatial variation in public service provision are considered bear the assumption that public services are smoothly "diffused" over exogenous jurisdictional boundaries, rather than being provided at just a few discrete locations in the urban area. One of the few theoretical models that have included a consideration of the impact of public facility location on the interdependent locational choices of firms and households is that of Thisse and Wildasin (1992). Depending on the value of critical parameters, their model concludes that firms can locate at the extreme ends of the urban area, together, or they might choose intermediate locations. Our empirical model analyzes the location of firms in the electronic industry in the Greater Madrid area. The model can be interpreted as a Tiebout type model, in terms of the resolution of the basic preference-revelation problem in the theory of public goods.

Empirical research on intra-metropolitan firm location has been hindered by the lack of disaggregated data and appropriate models. This paper moves forward in both aspects. Concerning the econometric models, a methodological framework that exploits all the intra-urban variation in spatial characteristics, and that has not previously been implemented in analyzing the spatial aspects of local public goods, is introduced. Our framework leverages information on the exact locations of establishments and public goods, and is not constrained by arbitrary scale-effects imposed by using data summaries in administrative boundaries. Traditionally there have been conceptual problems associated with the use of expenditure data as a measure of public services because variation in expenditure does not necessarily imply variation in the quality of the services. The spatial dimension of public services show that data constraints can be partially overcome by the use of GIS techniques in the modeling methodology. If the data are available at the site specific individual level models with a high level of precision can be estimated.

Much of the industrial district and agglomeration economies literature assumes that firms benefit from being located near other businesses. However, observed co-location is not enough to conclude that localization economies are driving the observed pattern; co-location may occur without linkages or interaction between proximate firms (Gordon and McCann, 2005; Torre and Rallet, 2005; Wai-chung et al., 2006). The modeling approach utilized identifies to what degree clustering is due to features of the environment (such as access to roads or other public amenities) or to direct interaction among firms (such that firms choose to locate near each other).

Location factors depend on the type of industry and on the size of the firm. Whether a firm's size affects its location preferences or site characteristics is checked by dividing the sample into large, medium and small firms. In the literature some papers find localization economies associated with smaller establishments (Figueiredo et al., 2009) while others report a relationship resembling an inverted "U" (Sweeney and Feser, 1998), and many studies suggest the size-agglomeration relationship depends on characteristics of an industry or methodology (Lafourcade and Mion, 2007; Duranton and Overman, 2008). The analysis is focused on the electronics industry for several reasons. First, by concentrating on a single narrowly defined industry the problem of unobserved inhomogeneity is reduced. Second, regional growth theories predict that clustering will be particularly strong among high tech or knowledge intensive industries. And third, it is a strategic industry in the region with a high degree of interaction and local interconnectedness.

The following section briefly describes the proposed modelling framework. Section 3 introduces the theoretical model. Section 4 describes the data and the process utilised to construct the covariates, and section 5 discusses the main results obtained. Concluding remarks are presented in section 6.

2. THE MODELLING FRAMEWORK

Studies on intra-urban firm location have employed a wide variety of techniques to determine the main factors that affect location. In all approaches implemented so far, the unit of analysis is an administrative metropolitan area or county for the US, and region (NUTS 2 or 3) or municipality in Europe. Survey researchers, probabilistic choice models of site demand, and traditional bid rent or hedonic methodologies have been employed (see Arauzo et al. (2010) for a detail review on these approaches). The basic framework of econometric empirical studies has essentially remained unaltered over the past four decades with different versions of either discrete choice models or count data models. Hedonic or bid rent models involve regressing land or space rents[1] on a host of location characteristics to identify those attributes that determine location desirability.

Spatial point processes can be used directly, to model and analyze data which take the form of a spatial point pattern, such as firm locations derived from geo-referenced addresses. The spatial distribution of firms is then considered to have arisen from a stochastic point process, with the observed distribution of firms being a single realization. Below we work with Gibbs

[1] The main assumption is that property rents embody valuable information on what might attract or repel firms to or from certain locations.

(Markov) processes, X, that can be expressed as exponential family densities and allow for separate estimation of effect sizes on components of the trend (“first-order effects”) and specific representation of the interaction (“second-order effects”). Details on the Gibbs process formulation can be found in Møller and Waagepetersen (2007).

A general expression of an inhomogenous Gibbs point process density (with respect to Poisson unit density) is,

$$f(x) = a \prod_{y \subseteq x} b(y) \gamma(y) \tag{1}$$

with normalizing constant *a*, localized shift *b(y)* in density at location y, and interaction function $\gamma(y)$ measuring interaction among points (in pairs, triples, or higher orders), and the product is over all subsets of x. In Sweeney and Gómez-Antonio (2013) several alternative interaction specifications fitted to firm locations are evaluated. The Geyer saturation form of interaction yielded the best diagnostics and so we only test models of that form here. In that model, the interaction is specified as $\gamma^{min\{d, N_x\}}$, so that interaction is a function of the number of R close neighbors N_x, but is bounded above by the scalar *d* so that the contribution of any point to total interaction is restricted. Without that restriction the process simplifies to a Strauss process which is known to yield only "single ball" clusters when $\gamma > 1$ (Møller and Waagepetersen, 2007). With that restriction the resulting saturation interaction is capable of generating moderate positive interaction that is characteristic of firm location patterns[2].

The normalizing constant, *a*, makes it difficult to work directly with the density, and the model is made tractable by working instead with the conditional intensity function,

$$\lambda(u, X) = b(u) \gamma^{min\{d, N_X(u)\}} \tag{2}$$

The conditional intensity focuses on the localized probability of observing an event in the vicinity of *u* given the rest of the process *X*.

Estimation:

Because of the exponential form of the Gibbs model, standard software implementations for generalized linear (additive) models can be used to estimate parameters of the conditional intensity function. Expanding the right hand side in log-linear form, the model is decoupled into two components,

$$\lambda(u, x) = exp\{\varphi^T b(u) + \vartheta^T S(u, x)\} \tag{3}$$

On the right hand side, the first term is the spatial trend component and the second term is the interaction component. Canonical parameters $\theta = < \varphi, \vartheta >$ are the focus of estimation. Note that the trend component depends only on the spatial location *u*, and reflects spatial in-homogeneity that affects the location decision of firms. Two groups of covariates are

[2] In the spatial statistics literature, Gibbs models have typically been employed to model processes of inhibition between events, while other types of model such as Cox models or cluster Poisson models are supposedly more appropriate to model attraction (clusters). Nevertheless the types of cluster process that might be found in other fields such as biology or physics are very different to the cluster structure that might present the spatial distribution of firms. In those fields cluster structure is characterized by a large concentration of events very close to one another, while in the establishments' patterns we assume a moderate level of clusters; Gibbs models are consequently more suitable. Another reason to estimate a Gibbs model is that we also need to account for spatial inhomogeneity, which is not considered in Poisson cluster type models.

introduced in this part of the model, related either to firms' costs or profits (land prices, population density, distance from distribution hubs, distance from CBD, taxation variables), or to the distance from certain local public goods (public transportation infrastructures, geographical knowledge spillovers and cultural and recreational facilities).

Estimation of the canonical parameters is based on the Huang and Ogata (1999) maximum likelihood as encoded in the R package *Spatstat*[3]. That method also provides an estimate of the asymptotic variance-covariance matrix for the canonical parameters. Standard errors and significance tests for those parameters are based that covariance estimate. The interaction component contains two irregular parameters, the range of interaction, R, and the threshold capping interaction at d. Profile likelihood is used to estimate the irregular parameters[4].

Diagnostics:

In addition to familiar specification tests of individual canonical parameters or groups of parameters using the estimated covariance, the complexity introduced with separate trend and interaction components requires a more varied set of diagnostic and specification tools. Baddeley et al. (2005, 2012) have extended diagnostics designed for GLMs to the point process setting. That translation is made possible by defining (pseudo-) innovations and (pseudo-) residuals for spatial point processes; accomplished essentially by defining the difference in observed points and expected points in subspace balls of the study region. In practice those balls are defined as grids in the study region.

Diagnostic measures and tests can be defined to isolate the adequacy of fit in the trend, the interaction, or the overall model. All of the methods we use are available in the R software using the *Spatstat* package, and the articles Baddeley et al. (2005, 2012) contain details of the theory underlying the use of pseudo-residuals. The diagnostics for trend include lurking variable plots and contour plots of smoothed residuals. Lurking variable plots indicate the cumulative error, measured as exponential energy weights in this case[5], over the domain of a covariate in the trend term, $b(u)$. The plots contain the expected value under the "null" -- that the fitted trend is correct -- and point-wise confidence envelopes (2 s.d.). Smoothed residuals are used to identify areas in the overall fitted trend that either over or under predict; a well fitting model would have a spatially mixed pattern of slight under and over fitted akin to points around a trend line in a scatter plot.

The fit and validation of the interaction term is assessed using the QQ-plot and the G-compensator. The QQ-plot compares quantiles of the smoothed residuals to expected quantiles of residuals under the fitted model. If the interaction component is captured by the model then the observed and expected quantiles will be within the error bounds provided in the plots. The G-compensator is an analog of a score test (Baddeley et al., 2005). It compares a nearest-neighbor distribution measure, the G-function, of the observed data and the expected

[3] Baddeley and Turner (2005)

[4] When estimating the model with a Geyer interaction specification, two irregular parameters need to be estimated: the saturation threshold, and the interaction radius of influence. We use a small set of integer values for the saturation parameter (1 to 8) and interaction radius (0.2 to 3) and select the combination of these that maximizes the profile pseudo likelihood. In the model with a Strauss Hard Core interaction specification the irregular parameters to be estimated are the Hard Core distance and the interaction radius, while for the Area Interaction component the interaction radius must be estimated.

[5] See Stoyan and Grabarnik (1991).

value under the fitted model. Again, error bounds are provided and if the fitted interaction term is well-matched to the underlying process of the data, the resulting G-compensator measure will fall within the error bounds.

3. THE MODEL

Blair (1987) identifies two stages in selecting a site at which to locate. The first stage consists of selecting a broad region focusing on labor, taxes, climate, proximity to markets and other features that may have significant interregional variation, but are similar everywhere within the region. The second stage consists of analyzing locational factors that vary at the micro-geographic level of detail. Analyzing the relevant factors that affect this second stage decision is the focus of the paper.

The set of explanatory variables used when analyzing industrial location varies considerably across studies, even when the studies share the same econometric specification. The theoretical model developed by Erickson and Wasylenko (1980) is adapted and extended to capture the representative factors analyzed in the theoretical literature, and factors that we consider may be important in the location choice of the electronics industry in the greater Madrid area. Consumer demand for output and therefore firms' revenues are assumed not to vary with intra-metropolitan locations. Differences in intra-metropolitan location affect firms' profits only through variation in their costs between locations.

Firms are assumed to maximize profit under the following production function:

$$Q = f\,(N, K, L, AG, G) \tag{4}$$

where N, K, L are labor, capital, and land respectively, AG are agglomeration economies, and G are local public goods. The role of three local public good is emphasized: transportation, knowledge spillovers, and recreational or cultural public goods.

The intra-metropolitan location decision is one of cost minimization, and the cost function per unit of output can be represented as:

$$C = f\,(PL, PN, S, t, G, AG) \tag{5}$$

where PL and PN are the price of industrial land and the price of labor, respectively. Some of the effective input prices are not available for sub-regional or sub-metropolitan areas. Wage rates figures could not be included as they were not available for individual communities of metropolitan areas[6]. S is a vector of general site characteristics, such as distance from central business district (CBD), proximity to distribution hubs or access to markets and labor. Finally, t is the effective property tax rate.

[6] Even if computed, they would be more a measure of the employment mix in the area than a measure of the wage rate that a firm in a particular industry must pay (Charney, 1983). The same problem was present for other covariates such as population density, industrial land price, and crime. Industrial land prices were available disaggregated by street and block, but they are not geo-referenced and the process of geo-referencing them and preparing them would constitute a highly time consuming and laborious task that goes beyond the scope of this paper.

In accordance with the estimation framework described in section 2 covariates are grouped into two components: the trend including all covariates related to environmental characteristics; and the term that captures the interaction between the firms.

A) The trend component

One of the classic factors in industrial location literature is industrial land price (PL). The expected sign is negative; however, land price might capitalize the value of local public services, and thus affect location choices. High-tech firms might prefer to locate in areas where land price is high. Three covariates related to general site characteristics (S) were included. The covariate distance to CBD capturing the importance of proximity to the central city markets for labor and goods[7]. The variable population density evaluates the importance of urbanization economies, in terms of access to a larger labor and output market. A positive sign is expected for this covariate. Finally, the relationship between the location of distribution hubs and firms' location choices is tested.

Governments can attract business by manipulating fiscal policies (*t*) and public services (*G*) to provide a profitable environment for firms. In the US economy, state and local taxes have an important effect on business location, particularly within metropolitan areas where business property taxes can vary substantially between jurisdictions. Earlier studies were not conclusive on the effects of taxation on industrial location (Luger and Shetty, 1985; Buss, 2001). Carlton (1979, 1983) found a non-significant effect of tax levels on location decisions in the United States; Bartik (1985) found that taxation exerts a moderately negative effect on U.S. states location. Charney (1983) found that relatively high levels of property taxes represent a significant disincentive to relocating firms selecting sites in an urban area. On the contrary, Gabe and Bell (2004) have argued that high-tax locations remain attractive as long as they spend large sums of money on the provision of public goods and services. The effective property tax rate is measured by the per capita payable tax (CI_IBI), which also controls for the nominal value of the assets. Although this value is updated periodically there are large differences between municipalities. A significant relationship between taxation variables and industrial location is not expected in Spain, as main taxes are centralized in the national government.

Government services in the area are related to the consumption of local public goods which are specific to a given geographical area. To derive utility from the public good, one must live in the community in which it is provided, and utility decreases with distance. Three local public goods, where distance might be an important component in their consumption, are considered: transportation access, access to knowledge spillovers, and access to cultural and recreational facilities. Closeness to the public good is assumed to reveal preference for its consumption.

1. Access to transportation is a local public good that represents an amenity for both firms and workers. The effect of public transport infrastructures has been extensively studied and it

[7] It will allow the hypothesis of the mono-centric or poly-centric central business districts to be checked.

differs across manufacturing industries,[8] indicating that accessibility requirements may vary with technology and/or demand. Access to main roads and proximity to the airport captures the benefits that firms enjoy in terms of time and lower transportation cost due to congestion. On the other hand, proximity to public transportation systems such as the metro, or distance to local train stations are local public goods that can be considered as worker amenities. Traffic congestion represents one of the main negative externalities in big cities, with Madrid being no exception. Access to these transportation systems represents a reduction in travel time, and an increase in comfort, due to their safety and frequency of service.

2. The public good nature of knowledge arises from nonproprietary university research. The role of these institutions centers on the existence of knowledge or technology spillover effects and the extent to which their diffusion is facilitated by geographical proximity. If productivity effects arising out of area universities are "universal" – that is, if they are present on a metropolitan-wide basis without significantly varying across locations – then the exact location of universities within urban areas may not influence the intra-urban geography of high-tech companies. The importance of geographic proximity depends on the "transport" mechanism of the spillover effects (Jaffe, 1989). If the mechanism is journal publications, geographic links are unimportant, but if it is through informal interactions between research managers and university researchers or through research contracts between academia and industry, then geographic links matter. Some of the benefits of proximity have to do with the lower cost of acquiring technological spillovers; with facilitating the diffusion of novel knowledge arising from university research; or with providing for the continuing education of employees, with access to faculty consultants and student interns (Bania et al., 1992). Of the Marshallian externalities that motivate the literature on agglomeration economies, knowledge spillovers have proven to be the hardest to verify empirically. Gibbs models constitute an appropriate tool to detect this type of Knowledge spillovers. There is mounting evidence for the attenuation of human capital spillovers at small spatial scales (Carlino et al., 2011; Wang et al., 2012).

3. Finally, the literature establishes that high-tech firms directly take into account the locational preferences of their potential professional employees, despite other locational influences arising from organizational purposes. Firms employing skilled professionals are supposed to be amenity oriented when selecting sites within metropolitan areas. In the survey literature, executives consistently rank *“quality of life”* as a top location factor, and high-tech firms are the most likely to consider residential amenities in their location decision (Goettlib, 1995). In the inter-area analysis Markusen et al. (1986), "area amenities" appear to be very important determinants of metropolitan high-tech employment. According to Florida (2005) it is the quality of places that attract creative people, and their presence attracts high-tech and cultural industries. This phenomenon “Cultural and recreational opportunities” is considered a local public good and three indicators are included in the model: a district´s green area extent; distance from the nearest movie theater[9]; and violent crime, considered as one disamenity that influences firm location when evaluated at the worksite itself. However, these *“quality of life”* amenities could be hypothesizes that are not likely to be a decisive location factor to firms

[8] Baudewyns et al. (2000), Holl (2004), Arauzo (2005), Alañón et al. (2007), Smith and Florida (1994), Luker (1998), Coughlin and Segev (2000), List (2001).

[9] Theaters are not confined to within the agglomeration, and are not equally scattered within it, so these might act as a good proxy for the cultural activity of the area.

selecting a site within a metropolitan area. These amenities could be relevant during the first stage of location, in choosing which urban area, city, or municipality to locate. But once this decision has been made these services might not be relevant at the intra-metropolitan scale. There can be a mismatch between the location of one's job and the location of one's residence as a firm may derive all possible benefits in the form of spillovers. Residents commute great distances because their residence environment offers amenities they consider to be more important than commuting costs. The Gibbs model approach implemented here is very suitable for testing this hypothesis.

B) Interaction component

One of the main strengths of this approach is that it allows testing for the importance of different types of spillovers in the industry. Agglomeration economies are the most studied determinant of industrial location. Empirical evidence of the positive effect that agglomeration economies have on location decisions are numerous: to quote just a sample, Luger and Shetty (1985), Hansen (1987), Coughlin et al. (1991), Friedman et al. (1992), Woodward (1992), Wu (1999), Coughlin and Segev (2000), Guimaraes et al. (2000), List (2001), Figueiredo et al. (2002), Holl (2004), Arauzo (2005), and Sweeney and Gómez-Antonio (2013). This approach allows us to measure more accurately the existence of agglomeration economies. It disentangles the effect of externalities determined by proximity to labor pools and knowledge spillovers, in the trend component, from the effect of other sources of spillovers captured in the interaction component. It naturally considers in the interaction component a measure of the advantages firms obtain from locating close to firms within the same industry, while the advantages firms obtain from locating close to firms of other industries (urbanization economies) are captured in the trend part of the model.

Once we have defined the covariates affecting the trend and the interaction component , the conditional intensity we estimate is the following:

$$Ln(\lambda(u,x)) = \alpha_0 + \alpha_1 PL + \alpha_{2,3} S + \alpha_4 AG + \alpha_5 t + \alpha_{6\text{-}12} GTransp + \alpha_{13} Gknow + \alpha_{14\text{-}16} Glife + \alpha_{17} Interaction \tag{6}$$

4. Data

The data used in the analysis refer to the year 2002. The geographical coordinates and the number of employees in each industrial electronics establishment are supplied by the Statistic Institute of the Region of Madrid, which also supplied the maps required for the analysis. The model is applied to the industry 32, according with the classification of the National System of Economic Activities (CNAE), "*Manufacture of electronic equipment, manufacturing equipment and radio, television and communication devices*", in the region of Madrid.

Point pattern analysis requires information on point locations of objects analyzed and the boundary that defines the subset of all possible point observations. In our case the observation window is a map of the Region of Madrid stored in the form of a GIS *shapefile*. For computational purposes, the observation window was simplified to the area with the greatest density of establishments surrounding Madrid's capital. Large areas in which no establishments are located were eliminated to reduce the computational burden of the model fitting.

There is no agreement between countries and industries to determine whether there are marked differences in location patterns between small and large plants. To determine possible differences in firms' location patterns depending on their size; the sample was divided according to the following classes: 1-4 employees, 5-9 employees, and 20 or more employees. The necessary maps were provided by the Madrid Regional Statistics Institute through the service *NOMECALLES*, in a *shapefile* format to compute the distance covariates. A decay effect of 0.25 was introduced to give more weight to firms closer to the facility. The variables included are distance to center; to different types of roads (freeways, ring roads, and radial roads); to subway and train stations; to airport; to distribution hubs; to technical universities; to theaters; and to green areas.

Two different procedures were followed in order to assign values to covariates that are not geo-referenced to specific sites. For the covariate population density, each of the polygons (districts[10] or municipalities) was randomly assigned as many points as the value of the covariate. Then a kernel smoother was implemented to make the covariate continuous in space. However, when the covariate was not divided by area, this procedure would be biasing the results in favor of the smaller districts/municipalities: two polygons with the same covariate value but different areas would present different intensities when they should be equal. To avoid this bias, a different procedure was followed for the covariates industrial land price, property tax, and surface of green area. A grid of 150 by 150 points (points are 300 meters apart) was overlaid and then to each intersection point in this grid, the value that the covariate takes in the polygon in which the point falls was assigned. The kernel smoother was then implemented to make the covariate continuous in space.

The crime rate was provided by the GESI (International Security Study Group), belonging to the Spanish Home Office. The number of arrests made by the National Police (*Policía nacional y Guardia Civil*) per 10,000 inhabitants was utilized. Since this rate was not available for Madrid municipal districts, in order to obtain a homogeneous series, Madrid's city data was obtained based on local police arrests by district.

The covariate industrial land price (measured in Euros per square meter) is the minimum value for the purposes of calculating the tax base of the Inheritance and Gift Tax. This value is intended to estimate the real value of the property when industrial land is transferred. The estimation method analyzes the real estate market on the date of accrual of the tax and calculates mean values for different areas from a sufficient number of "witnesses", consisting of properties of the same type and characteristics to the land being transferred[11].

---

[10] Due to the extent of the Municipality of Madrid we used the data of the 21 districts that is the lowest level of aggregation, while the rest of the municipalities are not divided by district.

[11] In some of Madrid´s districts there is no industrial land use. We thus introduced the price per km^2 of commercial lot as a robustness check, and results were confirmed.

The districts' population figures were obtained from the database of the Town Hall in Madrid, and the municipalities' populations from the Statistics Institute of the Region of Madrid. The area of the municipalities is provided by the *Statistical Yearbook of the Region of Madrid, 1985 to 2011: Environmental Conditions*. District areas were provided by the *Statistical Yearbook of the City of Madrid*.

Finally, the Green Urban Areas and Sport and Leisure Facilities were provided by CORINE Land Cover Project developed by the European Environment Agency included in the Land Core Monitoring System of the Global Monitoring for Environment and Security[12]. A map of the municipalities and district boundaries was overlaid in order to assign values to these polygons. Then a grid of 150 by 150 points was overlaid as explained in previous paragraphs.

5. RESULTS

To assess the validity of the results and prove that the model fits the observed establishment pattern, several specification diagnostics presented in the annex are briefly discussed below. In Figures 1-3 the lurking variable plots are shown for the subset of significant covariates in the models. As observed, the true spatial trend can be approximated by the specified trend. The empirical plot is close to its expected value assuming that the model is correct. The graphs do not exceed the point-wise two standard-deviation error limits calculated for the inhomogeneous Poisson process (dotted lines). When plotting the residuals of the model fitted to all the establishments' sample, the empirical line exceeded the envelopes for many of the covariates, illustrating the need to fit different models depending on establishments' size[13]. Additionally, as is shown in Figure 4, the smoothed residuals diagnostic presents a flat surface with small deviations from zero in all the models. Figures 5 and 6 show the QQ-plot and the G-compensator diagnostic, respectively. As is shown in the QQ-plot, the empirical distribution of the smoothed residuals lies inside the envelopes (estimated by Monte Carlo samplings) of the expected empirical quantiles obtained from simulations of the model with a Geyer interaction specification. In the G-compensator diagnostic, shown in Figure 6, the standardized residuals lie inside the envelopes showing positive values, suggesting that the data are slightly more clustered than the model. Overall the presented diagnostics indicate that the model with the Geyer saturation interaction term correctly captures the dependence on the covariates and the interaction between the establishments.

Estimated coefficients obtained from estimating Equation 6 are shown in Table 1. Results differ between those in which the model is estimated for the whole sample and those obtained when firm size is considered in the analysis.

(Table 1 around here)

First, small firms, due to the risk associated with their size and their inability to internalize certain necessary stages of production, tend to locate in dense areas, close to markets and support services. The covariate distance to center is significant at the 95% confidence level

---

[12] Land use 1.4.1 and 1.4.2 respectively.

[13] The complete set of lurking variable plots is available from the authors. The full set is not provided in the paper due to space constraints and to maintain visual clarity in the published figure.

only in the small firm's model, but with a positive sign. This indicates that in the area of study a polycentric structure is more prevalent than a monocentric one, reflecting the overwhelmingly decentralized nature of the industry (Shukla and Waddell, 1991). This result contradicts those obtained by Arauzo and Viladecans (2009) who show that Spanish manufacturing establishments in high-tech industries prefer to locate close to the center of the metropolitan area. It also challenges the results of Alañón et al. (2007), Arauzo (2005), Figueiredo et al. (2002), and Guimaraes et al. (1998, 2000). These studies are interurban area/region analyses where different urban areas or cities are considered, so the results are interpreted in terms of the average urban area. The results in intra-industrial location analyses are inconclusive. Wu (1999) determines that the traditional center has seen a declining attraction for new firms in the Metropolitan area of Guangzhou in China. Shukla and Waddell (1991) examine firm location in the context of poly-centricity, finding that the distance to the CBD (Dallas CBD) cannot provide a singularly powerful explanation of firm location, with the exception of the finance, insurance and real estate industries. On the other hand, Hansen (1987) for the state of Sao Paolo finds a negative sign for the distance to the CBD.

Second, the location decisions of firms in the electronics industry are dominated by site costs. For the small size sample, industrial land price is only significant at 95% level indicating that the importance of benefits from urbanization economies could dominate the price effect. As firm size increases, however, industrial land price becomes significant at 99% level in location choice. The analysis by Figueiredo et al. (2002) of the whole manufacture for the Portuguese districts, or that of Cheng and Stough (2006) of the location of Japanese FDI plants in Chinese cities obtained similar results.

Third, there is not significant evidence of a higher concentration of firms surrounding the distribution hubs in the subsets of data defined by employment size class. Firms do not reveal any preference for being in close proximity to a distribution hub, probably because the electronics industry do not use these logistics centers to ship their products or do not require large logistics infrastructures.

Fourth, the fiscal variable is only significant at the 90% confidence for the small firms subsample. The main local tax in Spain is the property tax (IBI), representing 50% of the local taxes and varying between municipalities. However, it represents a very small percentage of the tax burden faced by establishments.

Fifth, regarding the preference for local public goods, the results support the following conclusions.

a) Transportation costs are much less a consideration in location decisions in the electronics industry. At the 99% level only ring roads are statistically significant for large firm's model. The almost complete concentration of the telecommunication equipment manufacturing industry in Madrid probably contributed to the development of large establishments for which a location on the periphery of the city and near to the ring road M30 is more attractive. The ring road M40 presented the opposite expected sign. This result could be a consequence of the shape of this particular road, which extends far beyond the west surrounding the metropolitan area. The location of large firms is to some extent conditioned by land zoning restrictions, with most firms located in the eastern and northern municipal districts, and to a lesser extent in the southern districts. Ground transportation is expected to play a more significant role in industries with heavy or bulky products – those industries in which transportation costs are a large percentage of the total product cost – than in industries whose products are lighter and easier to transport (Smith and Florida 1994, McCann, 2011).

Firms have preferences for sites that are in proximity to train stations or metro transportation systems. Even large firms that locate in sites far from the city which might not be covered by the metro network are always in proximity of "local train stations". There is evidence of a significant preference for locations in proximity to Madrid International Airport and this is particularly the case for larger firms. Strauss-Kahn and Vives (2005) find the existence of airport facilities to be a significant location factor in analyzing the location of headquarters in US counties. Shukla and Waddell (1991) also find a significant effect in distance to airport in the analysis of Dallas urban area for the whole manufacturing industry. An analysis by Ihlanfeldt and Raper (1990) of the intra-urban location of new office firms finds that distance to airport is significant for new branch offices but not for new independent offices.

b) One of the most encouraging results is what we take as evidence of the existence of geographical knowledge spillovers. Larger firms present a higher probability of locating in the surroundings of technical universities. This result demonstrates the utility and relevance of the econometric approach we use to detect spillover effects. Papers that use administrative areas as units of observation cannot detect the existence of spillovers when they emerge inside the administrative boundaries. One of the reasons that proximity might be important is that in order to transfer new scientific findings into marketable products, face-to-face interaction and hands-on participation may be required (see Audretsch and Feldman, 1996; Audretsch and Stephan, 1996; Zucker et al., 1998; Gittelman, 2003). A number of commercial innovation centers such as Silicon Valley near San Francisco and Route 128 in Boston may respectively owe their status to their close ties to such prominent research institutions as Stanford and M.I.T. (Dorfman, 1983). The concentration of R&D firms in Chatsworth and Irvine in Greater Los Angeles may also be associated with benefits stemming from the California Institute of Technology and the University of California at Irvine campuses (Jürgen et al., 2004). Audretsch and Lehmann (2005) find that the number of knowledge-based start-ups clustered around German universities "is positively influenced by the knowledge output of the respective university and the innovative capacity of the region." In contrast to these findings, Howells (1984) reports that proximity to universities is not among the top considerations in the location decisions of pharmaceutical R&D in Britain. Wang et al. (2006, 2012) conclude that at the first phase of high and new technology industry clusters often happens close to universities named as the first "Cambridge phenomenon".

c) Cultural and recreational opportunities are not an important location factors in the electronic industry for selecting a site within a metropolitan area. These sorts of amenities are hard to measure because amenities and agglomeration are highly spatially correlated. It is difficult to distinguish whether high-tech location behavior is truly agglomerative or amenity oriented. Only large firms present a higher probability of locating in proximity to a theater. The green area surface in the district does not significantly differ from zero in any of the models. These results determine that in intra-metropolitan analysis we should be prudent about the relationship between 'quality of life' and the location of high-tech industries.[14] The high rate of commuting that characterizes the greater Madrid area determines that a firm may derive all possible benefits from amenities in the form of spillovers. Individuals can live in a jurisdiction with pleasant surroundings while working in an area with none of these characteristics. The covariate crime index that captures disamenities in the zone is always significantly different from zero but with an unexpected positive sign. Rosen (1979) cautions that crime may present an error-ridden proxy of exposure to crime, as different neighborhoods within heterogeneous cities may be subject to wide variation in crime rates. If crime figures were available at the neighborhood or at the street and block level of aggregation, the model would be able to check the validity of this hypothesis.

Sixth, the interaction component is always significant. The estimated parameter $\gamma>1$ indicates that the probability of observing a firm is higher if there is another firm in its proximity. Once the covariates that affect the first order intensity of the process are taken into account, the interaction component is significantly different from cero at 99% confidence level in all the estimated models. One of the explanations for the interaction result relates to geographical and professional proximity, as well as the similar origin of entrepreneurs, either from Movistar (former telefonica) or from the Polytechnic University, which encouraged the development of stable outsourcing relationships, prompting inter-firm collaboration as a means of minimizing capital risk (see Rama el al. 2003). The interaction result is in line with previous studies of the industry using different methodologies. Suarez-Villa and Rama (1996) and Rama and Calatrava (2002) determined a marked clustering of producers in the region, embedded in an intense web of subcontracting relationships, with the majority of establishments in Madrid obtaining a significant proportion of their production inputs, such as raw materials, parts and equipment, locally. López Bayón (2001) and Rama et al. (2003) confirm strong regional links between electronics firms in Spain and Madrid. Holl and Rama (2009) find that co-location implies strong interactions between firms in the electronics industry in Madrid. They determine that subcontracting linkages are among the most localized relations, suggesting that they involve close contacts for which proximity is desirable.

It is recognized in the literature that the Madrid electronics industry fulfills the characteristics required to be considered an industrial technological district as described by Park (1996), where interaction between the firms is bidirectional and non-hierarchical. Rama et al. (2003) found a high incidence of two-way subcontracting among Madrid's electronics producers, suggesting the existence of a complex, nonhierarchical inter-firm organization of production in Madrid's electronics district. Although in the 1980s the industry started with a hub-and-spoke model of production, where larger firms presented a unidirectional means of interaction by outsourcing parts of the production process to small and medium sized firms, the structure of the industry at this time is more prone to being considered as an industrial technological district. The fact that the interaction parameter is very similar for all size models might confirm this characteristic. In any case, this aspect is left for future analysis of hierarchical interaction between the firms.

[14] In the literature this covariate is proxied by the human capital level of the residents, the existence of good local schools, toxic emissions, landfill waste, per capita recreation expenditure, acreage of state parks, density of amusements employees, per capita local public expenditure, GDP per capita, distance to malls, and distance to sport and leisure facilities. None of them, however, were found to be statistically significant.

6. CONCLUSIONS

This paper improves the approaches used to date in measuring geographic concentration. Our improvements allow the disentanglement of first and second order effects, allowing a deeper understanding and quantification of different types of spillovers effects. Our results confirm the validity of the employed methodology for the analysis of intra-urban industrial location.

The intra-metropolitan distribution of electronics industry mainly follows a dichotomous pattern. Different covariates play different roles in explaining the trend component for each firm size class. The location patterns of small firms are mainly linked to current population and job distribution, but those of large firms are less affected by these variables.

In the greater Madrid area a polycentric structure is more prevalent than a monocentric one, where location decisions of firms in the electronics industry are dominated by site costs, and neither distribution hubs nor fiscal variables are factors considered by firms in deciding their location site. Regarding the local public goods analyzed, transportation costs are much less a consideration in location decisions in the electronics industry, but firms show a preference for proximity to train and metro stations, or the airport.

The results obtained challenge some of the outcomes of the intra-urban industrial location literature. The existence of geographical knowledge spillovers is confirmed for large establishments; amenities are not likely to be important location factors within a metropolitan area; and crime rate results are unexpected. All these results emphasize the need for different approaches in order to identify the factors that affect intra-urban location. Spatial aggregation compromises result in studies of business location. Intra-urban location decision processes might be rather different to inter-city or inter-regional location decisions. Firms consider different factors of localization at different scales, so results should not be compared between different spatial units of observation.

Interaction is significant in all models indicating the presence of sources of spillovers even after other sources of clustering are taken into account. A very particular form of spatial spillover is at work that could help to narrow the range of candidate theories to be tested. This type of interaction provides an interesting future research avenue.

The results of this paper suggest that Gibbs models constitute a fruitful approach to empirically explore the sources behind cluster configurations of firms at a variety of spatial scales. Gibbs models are sufficiently flexible to allow for multiple mechanisms, both for the purposes of identification and to avoid confounding results.

References:

Alañón, A., Arauzo, J.M., Myro, R., (2007) Accessibility, Agglomeration and Location, in Arauzo J.M., Manjón, M., (eds.), *Entrepreneurship, Industrial Location and Economic Growth*, Chentelham: Edward Elgar, 247-267.

Arauzo, J.M., (2005) Determinants of Industrial Location. An Application for Catalan Municipalities, *Papers in Regional Science*, 84, 105-120.

Arauzo, J.M., Liviano-Solis, D., Manjón, M., (2010) Empirical studies in industrial location: an assessment of their methods and results, *Journal of Regional Science*, 50, 3, 685-711.

Arauzo, J.M., Viladecans. E., (2009) Industrial Location at the Intra-metropolitan Level: The Role of Agglomeration Economies, *Regional Studies*, 43, 545-558.

Audretsch, D. B., Feldman, M.P., (1996) R&D spillovers and the geography of innovation and production. *American Economic Review*, 86, 630-640.

Audretsch, D., Lehmann, E., (2005) Does the Knowledge Spillover Theory of Entrepreneurship hold for regions? *Research Policy*, 34, 1191-1202.

Audretsch, D., Stephan, P., (1996) Company Scientist Locational Links: The Case of Biotechnology, *American Economic Review* 86, 3, 641-652.

Baddeley, A.J., Rubak, E., Møller, J., (2012) Score, pseudo-score and residual diagnostics for Goodness of fit of spatial point process models, *Statistical Science,* 26, 4, 613-646.

Baddeley, A.J., Turner, R., (2005) Spatstat: an R package for analyzing spatial point patterns, *Journal of Statistical Software*, 12 (6), 1–42.

Baddeley, A.J., Turner, R., Møller, J., Hazelton, M., (2005) Residuals for spatial point processes using the Papangelou conditional intensity, *Journal of the Royal Statistic Society, series B*, 67, 617-666.

Bania, N., Calkins, L.N., Dalenberg, D.R., (1992) The Effects of Regional Science and Technology Policy on the Geographic Distribution of Industrial R&D Labs, *Journal of Regional Science*, 32, 2, 209-228.

Bartik, T. J., (1985) Business Location Decisions in the U.S.: Estimates of the Effects of Unionization, Taxes, and Other Characteristics of States," *Journal of Business and Economic Statistics*, 3, 14-22.

Baudewyns, D., Sekkat, K., Ben-Ayad, M., (2000) Infrastructure Publique et Localisation des Entreprises `a Bruxelles et en Wallonie, in M. Beine and F. Docquier (eds.), *Convergence des regions: cas des regions belges*. De Boeck, pp. 281-305.

Blair J.P., (1987) Major Factors in industrial location: A review, *Economic Development Quarterly*, 1, 1, 72-85.

Buss, T.F., (2001) The Effect of State Tax Incentives on Economic Growth and Firm Location Decisions: An Overview of the Literature, *Economic Development Quarterly*, 15, 90-105.

Carlino, G.A., Carr, J.K., Hunt, R.M., Smith, T., (2011) The agglomeration of R&D labs, *Working Paper Research Department, Federal Reserve Bank of Philadelphia*, 11-42.

Carlton, D.W., (1979) Why New Firms Locate Where They Do: An Econometric Model, in W. Wheaton (ed.), *Interregional Movements and Regional Growth*. Washington: The Urban Institute.

Carlton, D.W., (1983) The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables, *Review of Economics and Statistics*, 65, 440-449.

Charney, A.H., (1983) Intraurban manufacturing location decisions and Local tax differentials, *Journal of Urban Economics* 14, 184-205

Cheng, S., Stough, R.R., (2006) Location Decisions of Japanese New Manufacturing Plants in China: A Discrete-Choice Analysis, *Annals of Regional Science*, 40, 369-387.

Coughlin, C.C., Segev, E., (2000) Location Determinants of New Foreign-Owned Manufacturing Plants, *Journal of Regional Science*, 40, 323-351.

Coughlin, C.C., Terza, J.V., Arromdee, V., (1991) State Characteristics and the Location of Foreign Direct Investment within the United States, *The Review of Economics and Statistics*, 73, 675-683.

Dorfman, N. S., (1983) Route 128: The Development of a Regional High Technology Economy, *Research Policy*, 2, 12, 299-316.

Duranton, G., Overman, H., (2008) Exploring the Detailed Location Patterns of UK Manufacturing Industries Using Microgeographic Data, *Journal of Regional Science* 48, 213-243.

Erickson R., Wasylenko, M., (1980) Firm relocation and site selection in suburban municipalities, *Journal of Urban Economies*, 8, 69-85.

Figueiredo, O., Guimaraes, P., Woodward, D., (2002) Home-field Advantage: Location Decisions of Portuguese Entrepreneurs, *Journal of Urban Economics*, 52, 341-361.

Figueiredo O, Guimarães P, Woodward D (2009) Localization Economies and Establishment Size: Was Marshall Right After All? *Journal of Economic Geography* 9, 853-868.

Florida, R. (2005), The world is spiky, *The Atlantic Monthly,* 296, 3, 48–51.

Friedman, J., Gerlowski, D.A., Silberman, J., (1992) What Attracts Foreign Multinational Corporations? Evidence from Branch Plant Location in the United States, *Journal of Regional Science*, 32, 403-418.

Gabe, T., Kathleen P.B., (2004) Tradeoffs between Local Taxes and Government Spending as Determinants of Business Location, *Journal of Regional Science*, 44 (1), 21-41.

Gittelman, M., (2003) Does geography matter for science-based firms? Epistemic communities and the geography of research and patenting in biotechnology, *SSRN Electronic Paper Collection*, Stern School, New York University.

Goettlib. P. D., (1995) Residential amenities, firms location and Economic development, *Urban Studies*, 32, 9, 1413-1436.

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

Guimaraes, P., Figueiredo, O., Woodward, D., (2000)Agglomeration and the Location of Foreign Direct Investment in Portugal, *Journal of Urban Economics*, 47, 115–135.

Guimaraes, P., Rolfe, R.J., Woodward, D., (1998) Regional Incentives and Industrial Location in Puerto Rico, *International Regional Science Review*, 21, 119-138.

Hansen, E.R., (1987) Industrial Location Choice in Sao Paulo, Brazil: A Nested Logit Model, *Regional Science and Urban Economics*, 17, 89-108.

Holl, A., (2004) Start-Ups and Relocations: Manufacturing Plant Location in Portugal, *Papers in Regional Science*, 83, 649–668.

Holl, A., Rama, R., (2009) The spatial patterns of networks, hierarchies and subsidiaries, *European Planning Studies*, 17, 9, 1261-1281.

Howells, J. R. L., (1984) The Location of Research and Development: Some Observations and Evidence from Britain, *Regional Studies,* 18, 1, 13-29.

Huang, F., Ogata, Y., (1999) Improvements of the Maximum Pseudo-Likelihood estimators in various spatial statistical models, *Journal of Computational and Graphical Statistics*, 8, 3.

Ihlanfeldt, K.R., Raper, D., (1990) The intrametropolitan location of New office firms, *Land Economics*, 66, 2, 182-198.

Jaffe, A., (1989) Real effects of academic research, *The American Economic Review*, 79, 957-970.

Jürgen E., Gottschalk, S., Rammer, C., (2004) Location Decisions of Spin-offs from Public Research Institutions, *Industry and Innovation*, 11, 3, 207-223.

Lafourcade, M., Mion, G., (2007) Concentration, Agglomeration and the Size of Plants, *Regional Science and Urban Economics* 37, 46–68.

List, J. A., (2001) US County-Level Determinants of Inbound FDI: Evidence from a Two-Step Modified Count Data Model, *International Journal of Industrial Organization*, 19, 953-973.

López Bayón, S. (2001) Características de la subcontratación electrónica en España: evidencias empíricas, *Documento de Trabajo 246, Facultad de Ciencias Económicas* (Oviedo, Spain: Universidad de Oviedo).

Luger, M. I., Shetty, S., (1985) Determinants of Foreign Plant Start-ups in the United States: Lessons for Policymakers in the Southeast, *Vanderbilt Journal of Transnational Law*, 18, 223-245.

Luker, B., (1998) Foreign Investment in the Nonmetropolitan U.S. South and Midwest: A Case of Mimetic Location Behavior?, *International Regional Science Review*, 21, 163-184.

Markusen A, Hall, P., Glasmeier, A., (1986) *High tech America: The what, how, where, and why of the sunrise Industries*. Boston

McCann, P., (2011) *Urban and Regional Economics*, Routledge.

Møller, J., Waagepetersen, R.P., (2007) Modern statistics for spatial point processes (with discussion), *Scandinavian Journal of Statistics*, 34, 643 - 711.

Park, S.O., (1996) Networks and embeddedness in the dynamic types of new industrial districts, *Progress in Human Geography* 20, 476-493

Rama, R., Calatrava, A., (2002) The advantages of clustering: The case of Spanish electronics subcontractors, *International Journal of Technology Management*, 24 (7/8), 181-229.

Rama, R., Ferguson, D., Melero, A., (2003) Subcontracting networks in industrial districts: the electronics industries of Madrid, *Regional Studies,* 37, 1, 71-88.

Rosen, S, (1979) Wage-Based Indexes of the Quality of Life, in *Current Issues in Urban Economics* Eds. Mieskowski, P., Strassheim, M., (John Hopkins Press, Baltimore, MD)

Shukla, V., Waddell, P., (1991) Firm Location and Land use in Discrete Urban Space. A Study of the Spatial Structure of Dallas-Fort Worth, *Regional Science and Urban Economics* 21, 225-253.

Smith, D.F., Florida, R., (1994) Agglomeration and Industrial Location: An Econometric Analysis of Japanese-Affiliated Manufacturing Establishments in Automotive-Related Industries, *Journal of Urban Economics*, 36, 23-41.

Stoyan, D., Grabarnik, P., (1991) Second-order characteristics for stochastic structures connected with Gibbs point processes, *Mathematische Nachrichten* 151, 95-100.

Strauss-Kahn, V., Vives, X., (2005) Why and Where Do Headquarters Move? *Discussion Paper 5070, Centre for Economic Policy Research.*

Suarez-Villa, L., Rama, R., (1996) Outsourcing, R&D and the Pattern of Intra-metropolitan Location: The Electronics Industries of Madrid, *Urban Studies* (33), 1155-1197.

Sweeney S, Feser E., (1998) Plant Size and Clustering of Manufacturing Activity, *Geographical Analysis*, 30(1):45-64.

Sweeney, S., Gómez-Antonio, M., (2013) On the utility of Gibbs models for refined assessment of localization economies: An application to the Madrid electronics industry, Mimeo.

Thisse, J.F., Wildasin, D.E., (1992) Public facility location and urban spatial Structure Equilibrium and welfare analysis, *Journal of Public Economics* 48, 83-118.

Torre, A., Rallet, A., (2005) Proximity and Location, *Regional Studies*, 39 (1), 47–59.

Wai-Chung Yeung, H., Liu, W., Dicken, P., (2006) Transnational corporations and network effects of a local manufacturing cluster in mobile telecommunications equipment in China, *World Development*, 34 (3), 520–540.

Wang, Z., Zhao, J.Y., Liu, X., Mao, K.J., Xie, X.L., (2006) An analysis to the evolution law of high-tech industry in space and its factors, *Studies in Science of Science* 24, 2, 227-232.

Wang, Z., Liu, X., Mao, K.J. (2012) Industry cluster: spatial density and optimal scale, *Annals of Regional Science* 49, 719-731.

Woodward, D. P., (1992) Locational Determinants of Japanese Manufacturing Start-Ups in the United States, *Southern Economic Journal*, 58, 690-708.

Wu, F., (1999) Intrametropolitan FDI Firm Location in Guangzhou, China: A Poisson and Negative Binomial Analysis, *The Annals of Regional Science*, 33, 535-555.

Zucker, L., Darby, M., Brewer, B., (1998) Intellectual Human Capital and the Birth of the US Bio- technology Enterprise, *American Economic Review* 88(1), 290-306.

Table 1: Models Estimation

Variable	All	Small	Medium	Large
Intercept	1.681 (1.225)	1.187 (1.849)	-1.879 (1.962)	8.351*** (2.723)
Industrial Land	-0.001*** (0.000)	-0.001* (0.001)	-0.002*** (0.001)	-0.002** (0.001)
Population density	0.001 (0.000)	0.002*** (0.001)	-0.000 (0.001)	-0.001 (0.001)
Distribution Hubs	0.328 (0.268)	0.491 (0.561)	0.998* (0.510)	-0.087 (0.539)
Property Tax	-0.002 (0.001)	-0.004* (0.002)	-0.003 (0.002)	0.000 (0.002)
Distance to Center	0.102 (0.065)	0.254** (0.105)	-0.018 (0.104)	0.166 (0.149)
Distance to road R	0.005 (0.021)	0.064* (0.036)	-0.062 (0.040)	-0.012 (0.040)
Distance to road A	-0.043 (0.036)	-0.035 (0.060)	-0.074 (0.059)	0.004 (0.059)
Distance to road M30	-0.143** (0.059)	-0.150* (0.084)	-0.091 (0.085)	-0.472*** (0.146)
Distance to road M40	0.022 (0.041)	-0.111 (0.073)	0.039 (0.071)	0.288*** (0.106)
Distance to airport	-0.887*** (0.331)	-1.054 (0.687)	-0.881 (0.598)	-2.331*** (0.702)
Distance to Local Train	-0.873*** (0.292)	-0.747 (0.466)	-0.746* (0.400)	-2.135*** (0.534)
Distance to Subway	-1.242*** (0.307)	-3.323*** (0.552)	-1.858*** (0.465)	0.227 (0.607)
Distance to University	0.059 (0.422)	0.904 (0.644)	0.112 (0.584)	-1.548** (0.729)
Distance to Theater	-0.347 (0.331)	-0.704 (0.544)	0.085 (0.478)	-1.231** (0.571)
Crime Index	0.013*** (0.004)	0.017** (0.008)	0.016** (0.006)	0.017** (0.007)
% Green area	-0.051* (0.030)	-0.115 (0.126)	-0.038 (0.029)	-0.064 (0.043)
Interaction	0.539*** (0.054)	0.285*** (0.085)	0.444*** (0.063)	1.276*** (0.233)

Notes: [a] Standard Errors are reported in parentheses below the coeficients
***,**,* Indicate significance at the 1, 5, 10 percent levels, respectively

ANNEX: LIST OF FIGURES

Figure 1: Lurking Variable plot. Small establishments

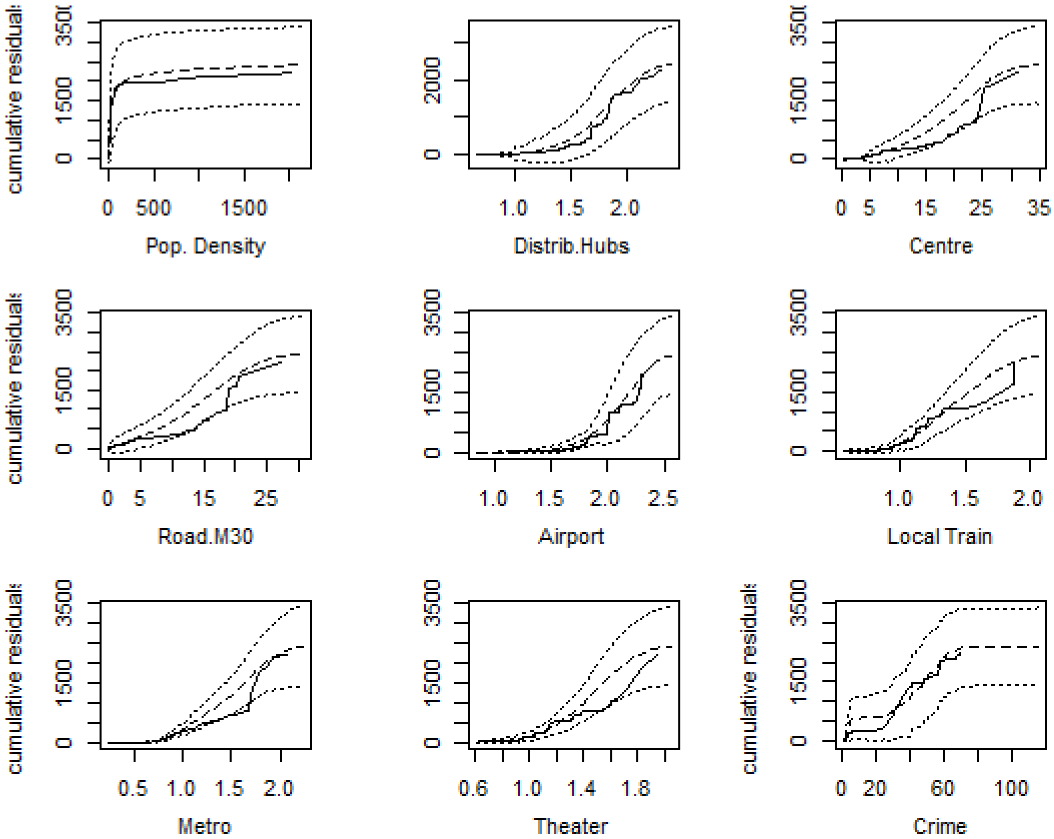

Figure 2: Lurking Variable plot. Medium establishments

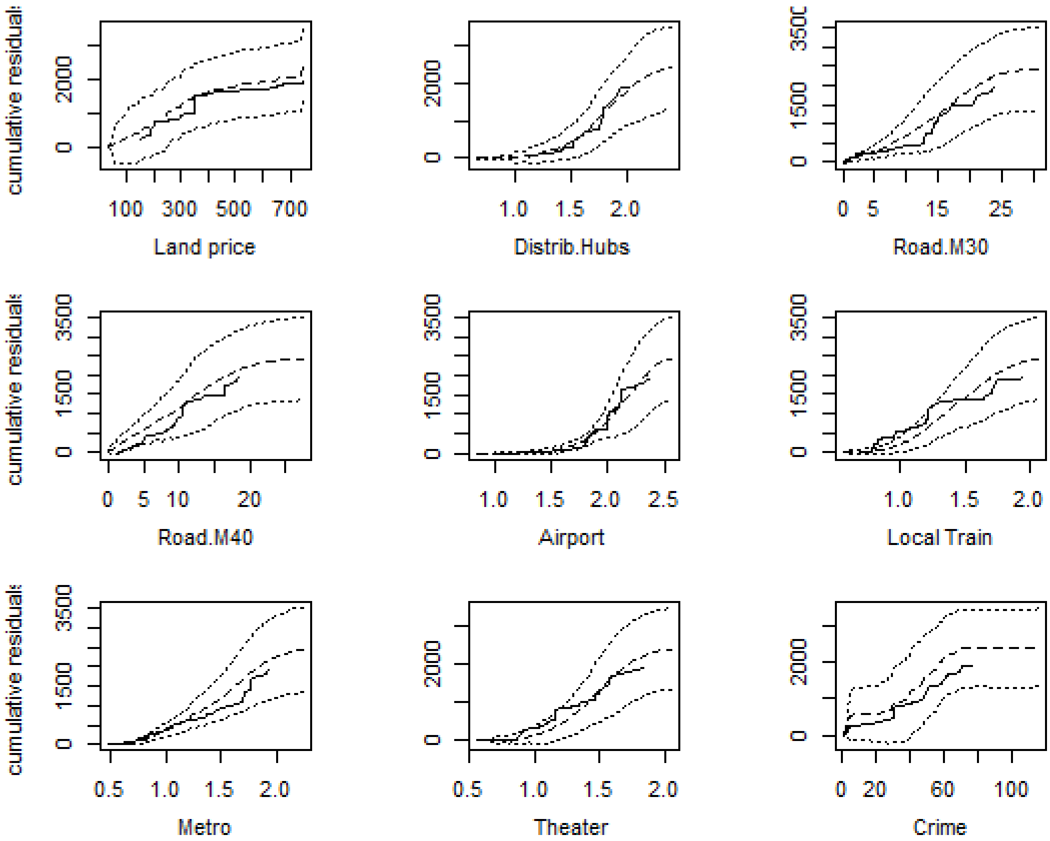

Figure 3: Lurking Variable plot. Large establishments

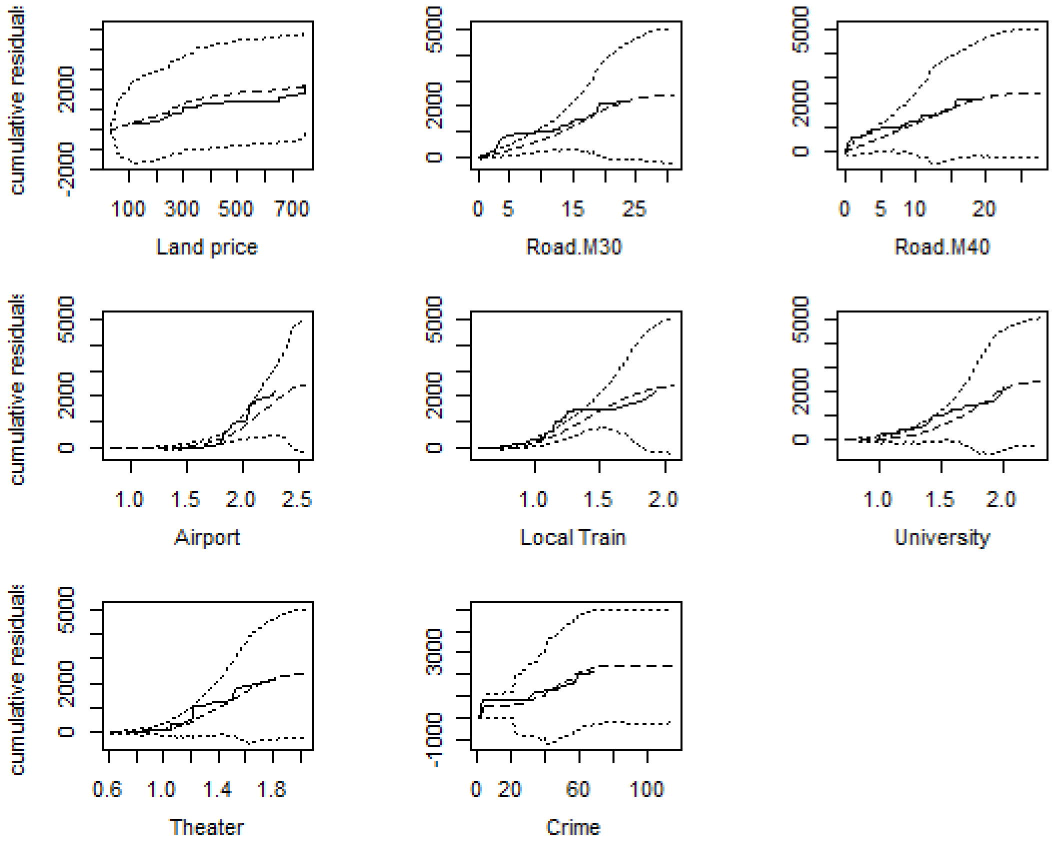

Figure 4: Smoothed Residuals

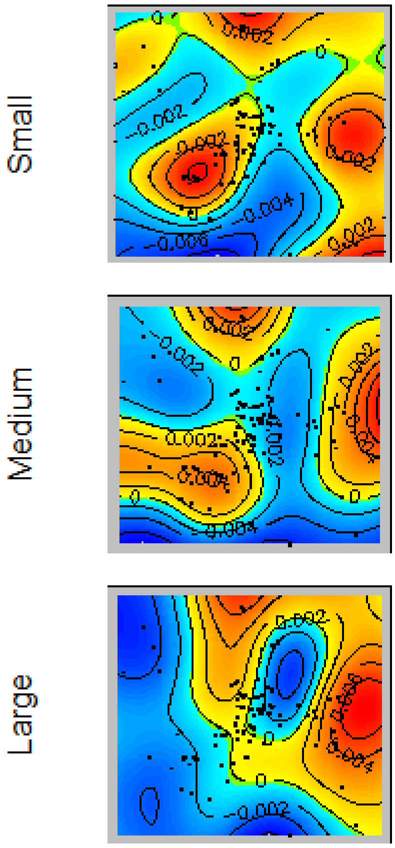

Figure 5: QQ-plot

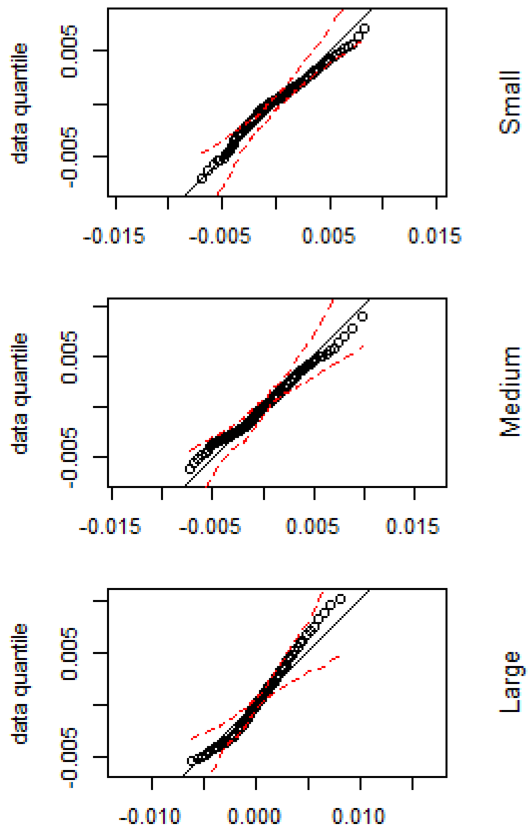

Figure 6: G-compensator