

# **The impact of accessibility by public transport on real estate values: a comparison between the cities of Rome and Santander**

Ruben Cordera<sup>a1</sup>  
Pierluigi Coppola<sup>b</sup>  
Luigi dell'Olio<sup>a</sup>  
Ángel Ibeas<sup>a</sup>

a Department of Transportation, University of Cantabria, Av. de Los Castros s/n, 39005 Santander, Cantabria, Spain

b Department of Civil Engineering, “Tor Vergata” University of Rome, Via del Politecnico 1, 00133 Roma, Italy

## **ABSTRACT**

Empirical evidence supports the hypothesis that greater accessibility to opportunities can have a positive impact on real estate values. However, this capitalisation of the benefits of accessibility could vary between different study areas. This research estimates hedonic and spatial hedonic models in two urban areas to compare if differences can be found in the impact accessibility using public transport has on residential property values. The two study areas chosen for comparison are a medium sized city without any major mobility problems (Santander, Spain) and a very large city with major congestion problems (Rome, Italy). The estimated hedonic models have considered the possible presence of spatial effects, a common occurrence with real estate data which may generate dependency in the residuals of hedonic models. Accessibility has been measured using two types of indicators: relative and gravity based. The results confirm that accessibility was a positive factor on property prices in both cities, although in Santander this was only true using the relative indicator to the city centre. These results are relevant for supporting the introduction of value capture policies which provide finance for new projects to extend and encourage greater use of public transport.

## **KEYWORDS**

Accessibility; Hedonic price models; Public transport; Real estate prices; Spatial – econometric models

---

1 Corresponding Author. Tel (+34-942-20-15-66). Email: [corderar@unican.es](mailto:corderar@unican.es)

## 1. INTRODUCTION

The benefits generated by transport in terms of accessibility to different opportunities can be capitalised by property owners (Martínez, 2000). Real estate is a heterogeneous good whose market price is influenced not only by its structural characteristics but also by the physical characteristics of the surrounding environment and the level of accessibility. According to the Henry George theorem, under certain theoretical conditions the aggregated cost in goods and services will increase real estate rents by the same value (Arnott and Stiglitz, 1979). The benefits generated by transport equipment or infrastructure can therefore be capitalised on by property owners which opens the way for value capture policies leading to a more efficiently financed public transport service.

Value capture policies are currently being applied in various cities using different mechanisms like coordinated urban and public transport development (Hong Kong, San Francisco, Atlanta), the application of direct taxes on property owners who are directly benefitting from transport infrastructure (Portland, Washington) and services or taxes levied on new real estate developments (Chicago) (Fogarty and America, 2008). These policies have generally proved to be effective in improving the economic sustainability and financing of public transport.

The weight of the different factors influencing property prices can be estimated using a technique known as hedonic regression. This technique was formalised by Rosen (1974), although multiple empirical studies had been previously developed since the pioneering work of Court (1939). Hedonic regression models have proven to be useful for evaluating the weight of different factors on heterogeneous goods like real estate (Malpezzi, 2008). By establishing a linear regression for all the relevant attributes of an item, the contribution of each factor for increasing or decreasing different environmental characteristics can be evaluated. However, the observations in this type of model may be spatially dependent which, when corresponding to the dependent variable omitted in the hedonic function, will generate biased and inconsistent parameters (Anselin and Rey, 2014).

This article proposes the estimation of hedonic regression models and hedonic regression models which consider spatial dependence between observations using aggregated data obtained in Rome (Italy) and Santander (Spain). These models will estimate if accessibility using public transport is a factor with significant weight on residential property prices. Conventional hedonic models will be compared with spatial hedonic models to determine if they have a significantly improved fit to the data.

The results show how accessibility using public transport is a factor which may have a positive impact on property values in Rome, whereas in Santander a significant

relationship can only be found measuring accessibility relative to journey time from each zone to the city centre. Furthermore, while the real estate data from Rome showed a strong autocorrelation in the residuals of the models, this autocorrelation was weak in the case of Santander and the spatial hedonic models did not show a great improvement over the traditional hedonic models.

The following section will introduce a brief state of the art in the field of the relationship between property prices and accessibility. The methodology used in this research is summarised in section 3, concentrating mainly on linear regression and regression considering spatial dependency between observations. Section 4 provides a description of the two study areas and presents and discusses the results provided by the models. Finally, the main conclusions drawn in this research are summarised in section 5.

## **2. STATE OF THE ART: THE IMPACT OF ACCESSIBILITY ON REAL ESTATE VALUES**

Hedonic models have been applied in different case studies for evaluating the impact of several environmental characteristics on real estate prices. Boyle and Kiel (2001) summarised multiple study cases addressing environmental externalities (air quality, water quality and distance to undesirable land uses) and showed that the estimated models generally had the expected sign in the parameters and were also statistically significant. Debrezion et al. (2007) performed a meta-analysis with over 50 studies about the impacts of railway stations on nearby residential and commercial property values. In this case, the authors found that the impacts on commercial properties were produced above all over short distances, while over longer distances the dominant effect was on residential properties. Mohammad et al. (2013) also performed a meta-analysis, in this case based on 23 studies, about the impacts of railway stations on real estate values. The variations in the resulting estimations from the studies were due, among other factors, to the differences between the types of property, the transport mode being analysed, the distance to the station and the geographical location.

Generally, empirical studies addressing the impacts of transport on real estate have had positive results showing that increases in accessibility correlate with higher property values. However, different studies have found neutral effects (Senior, 2009) and even negative effects in some cases (Bowes and Ihlanfeldt, 2001), particularly in the field of accessibility to railway stations due to the negative externalities associated with these facilities. In this area, Chen et al. (1998) found that the distance to light rail transit stations had a positive effect on single-family home values of the city of Portland (Oregon). However, they also identified a negative effect on prices resulting from the proximity to the line. Ibeas et al. (2012) found evidence from disaggregated data of positive impacts particularly if accessibility was measured as the journey time to the

city centre in the study area of Santander (Spain). Wang et al. (2013) confirmed the importance of accessibility to transport on land prices in a study area in Travis County (USA) using different indicators such as the distance to the shopping centre and the Central Business District (CBD) or the distance to arterial roads. Dubé et al. (2013) estimated models to measure the influence of public transport by train in Montreal (Canada). The authors could determine that housing close to the new stations increased their average market values by 2.6%, a rent that was captured by increased city taxes generating various millions of dollars in extra income for the system.

In Barcelona (Spain), Marmolejo and González (2009) studied the influence of an environmental factor, such as noise, on real estate prices using Geographically Weighted Regression (GWR). The authors also considered the effect of travel time to workplaces in the model. The research estimated a 0.6% drop in real estate prices per each additional minute of travel time to work.

Another line of investigation estimated the impacts on real estate values derived from the opening of new Bus Rapid Transit (BRT) systems. Munoz-Raskin (2010) used data from the city of Bogotá (Colombia) and found that the housing located nearest to the bus stops had an average value which was almost 10% greater than housing located over 5 minutes walking distance away. Jojoa and Marmolejo (2013) also studied, among other factors, the effect of the distance to BRT stations on the real estate prices. The authors found a negative gradient according to the distance to the stations and not only a binary effect such as that obtained by Munoz-Raskin (2010). Cervero and Kang (2011) studied the impact of a new BRT in Seoul (South Korea) and found increased property values for housing located under 300 meters from the bus stops. These increases were almost 10% for residential land use and 25% for non-residential land use considering a smaller impact distance of 150 metres.

The importance of social and environmental settings in real estate prices has also been examined by several authors. Factors such as the presence of green areas (Jim and Chen, 2010; Panduro and Veie, 2013), or economic (François et al., 1996) and social characteristics (Colombo and Stanca, 2014) have proven to be significant factors in defining real estate prices.

Some authors have highlighted the importance of considering the effects of spatial dependence which commonly occurs in the cross sectional data from the real estate market. Armstrong and Rodríguez (2006) highlighted how the presence of spatial dependence in the data in the form of autocorrelation may imply bias, inconsistency or inefficiency in the parameters of the hedonic functions estimated using ordinary least squares (OLS). The authors detected the presence of spatial dependence in the data of their study area (Eastern Massachusetts, USA) which led them to apply a model with spatial autocorrelation in the dependent variable. This allowed them to estimate that the

properties located in the municipalities with train stations had an average price of 10% higher than the rest. On a similar line, Martínez and Viegas (2009) applied autoregressive models in the dependent variable to determine if access to public transport was capitalised in real estate prices after detecting the presence of spatial autocorrelation in the data. The results supported the hypothesis that accessibility to metro stations could significantly increase property values even after controlling the effects of spatial dependence. Efthymiou and Antoniou (2013) estimated hedonic models for residential properties in Athens (Greece) considering the effects of dependency between observations and spatial heterogeneity in the study area. These effects of spatial heterogeneity were considered by the authors using GWR (Fotheringham et al., 2002). In all cases the distance from transport infrastructure had an impact on property prices, which was positive in the cases of bus stops, metro stations and tram stops, but negative in the cases of railway stations, airports and ports.

Most studies can be stated to have found that accessibility to transport and to public transport may have a positive impact on housing prices. However, the estimations of the parameters have proven to be very variable, suggesting that comparative studies applying the same methodology in different areas may prove useful. In addition, researchers have mainly concentrated on measuring accessibility to transport facilities and services (stations and stops) whereas much less attention has been focussed on access to the end opportunities, which are what really give utility to the users as the transport services are a derived demand. The following sections will describe the estimation of whether or not these impacts of accessibility to opportunities exist in the cities of Santander and Rome and whether they have a similar magnitude in both study cases controlling on the possible presence of spatial dependence in the data to guarantee the reliability of the estimated parameters.

### 3. METHODOLOGY

The most frequently used methodology for estimating the impact of different factors on real estate is hedonic regression. This method is based on a linear regression model specified as follows:

$$y = X\beta + \varepsilon \quad (1)$$

Where  $y$  is the price or the advertised price of a property or a group of properties, normally specified logarithmically,  $X$  is a matrix holding information about the independent variables, such as the structural characteristics of the housing, variables related to transport and the environment etc,  $\beta$  is a vector of the parameters to be estimated and  $\varepsilon$  is a vector of independent and identically distributed errors (IID). In this study, the variables of interest contained in the  $X$  matrix will be the indicators of

accessibility, while the remaining variables will act as control variables.

An introduction to the main spatial econometric models developed in the literature can be found in LeSage and Pace (2009). The most widely known spatial econometric model used for addressing spatial dependence in the observations of a linear regression is the simultaneous autoregressive model (SAR). The SAR model assumes the presence of a process of dependence in the dependent variable. This model is specified as:

$$y = \rho W y + X \beta + \varepsilon \quad (2)$$

Where  $\rho$  is the parameter of spatial autocorrelation,  $W$  is a  $N \times N$  spatial weights matrix being  $N$  the number of observations. The remaining variables are identical to those present in (1). Another well – known spatial regression model is the Spatial Error Model (SEM) in which only a spatial correlation between the error terms is specified.

$$y = X \beta + u \quad (3)$$

$$u = \lambda W u + \varepsilon \quad (4)$$

Where  $\lambda$  is the parameter of spatial autocorrelation in the residuals  $u$ . The neighbourhood matrix  $W$  can be specified in a variety of ways and can also take into account whether the available data are of a zonal or point nature. The 4 most common types of neighbourhood are: queen type, tower type, predetermined number of near neighbours and maximum neighbourhood distance (Anselin and Rey, 2014). The queen type contiguity considers as neighbours those adjacent locations that share a border or a vertex with the given location, whereas a tower type contiguity considers as neighbours those observations that share a border with the reference location (Anselin, 1988). Lesage and Pace (2010a) studied the spatial dependence which was present in different neighbourhood matrices aiming to demonstrate how the influence of the  $W$  specification on the estimations of the parameters was minimal if they are correctly interpreted from the true partial derivatives (direct and indirect impacts together), if the number of near neighbours is similar in all the observations and if the model is correctly specified.

Spatial dependence in the residuals of the hedonic regression models applied to the real estate market can occur due to a variety of reasons. Firstly, due to the effect of spatial diffusion of property prices on nearby observations. Secondly, because of the absence of equilibrium between the supply and demand for housing in different market areas. Finally, the omission of, or the partial measurement of, a relevant variable with differential effects on the space being studied could also generate a problem of dependence in the residuals (Bitter et al., 2007). Spatial econometric models help to minimise these effects by explicitly addressing the spatial dependence between observations.

The concept of accessibility to opportunities can be measured using different indicators. Handy and Niemeier (1997) classified these indicators into three large groups: accumulated opportunities, gravity type and those based on random utility theory. Gravity type indicators are those that have found the most applications on a practical level. Their general formulation is as follows:

$$A_i = \sum_j f(E_j, C_{ij}) \quad (5)$$

Where  $A_i$  is accessibility to opportunities in a determined zone  $i$ ,  $E_j$  is a measure of the attraction of zone  $j$  (e.g. jobs, commercial opportunities, etc.) and  $C_{ij}$  is a measure of the cost of making a trip between zones  $i$  and  $j$ . Compared to the accumulated opportunities indicators, the gravity indicators are able to differentially weight the opportunities according to the journey cost. Whereas in the accumulated opportunities models the costs are only weighted in a binary way, taking a value of 1 or 0 according to whether the opportunities are found inside or outside the determined range (Koenig, 1980). Compared to the utility based indicators, the gravity indicators have a zonal nature which makes them more appropriate for research focussed on an intermediate scale.

This work will apply the gravity type indicators proposed by Cascetta (2009) and Coppola and Nuzzolo (2011). These authors differentiate the active accessibility of a zone defined as its capacity to reach the opportunities present in other zones, from the passive accessibility as the capacity of zone to be reached by the populations of other zones. These indicators take the following functional forms:

$$ACC\_ACT_i = \sum_j [\exp(\alpha_2 \cdot C_{ij}) \cdot E_j^{\alpha_1}] \quad (6)$$

$$ACC\_PAS_j = \sum_i [\exp(\alpha_4 \cdot C_{ij}) \cdot P_i^{\alpha_3}] \quad (7)$$

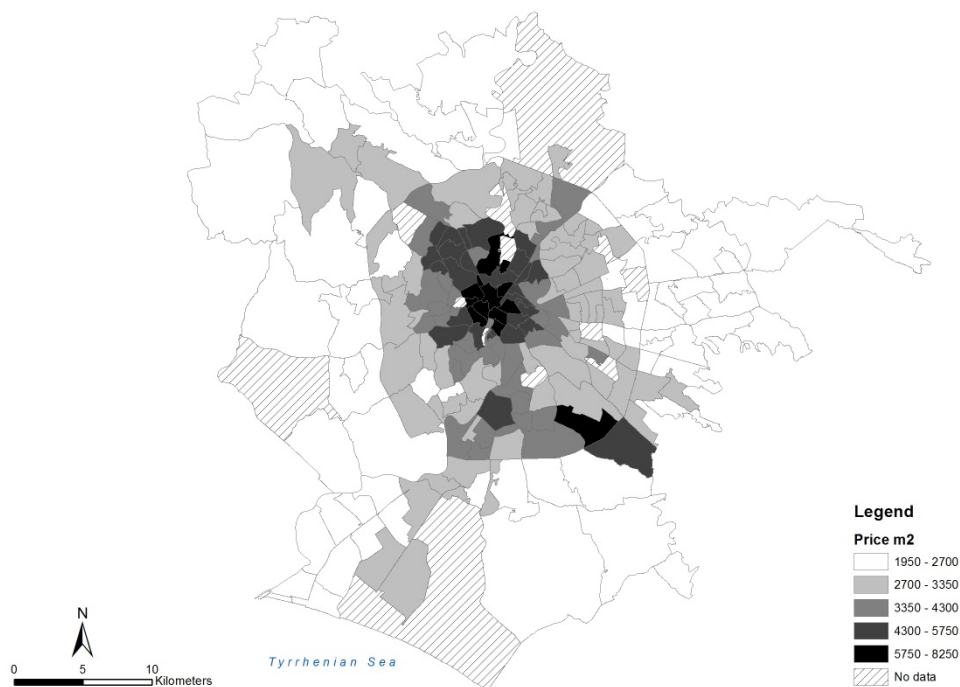
Where  $P_i$  is the population or the number of households present in zone  $i$ ,  $E_j$  and  $C_{ij}$  are equal to those present in equation (5) and  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$  and  $\alpha_4$  are parameters to be estimated. The parameters of the indicators can be calibrated using OLS taking logarithms to both sides of expressions (6) and (7). Furthermore, in areas with a clear monocentric nature, the travelling time to the city centre may be a good indicator of relative accessibility to opportunities such as jobs, leisure activities, shopping and other trip purposes. In this type of monocentric areas, the correlation between the gravity indicators and the accessibility indicators relative to the city centre can therefore be high which means they should not both be inserted into the hedonic function to avoid situations of high collinearity.

## 4. STUDY AREAS AND RESULTS

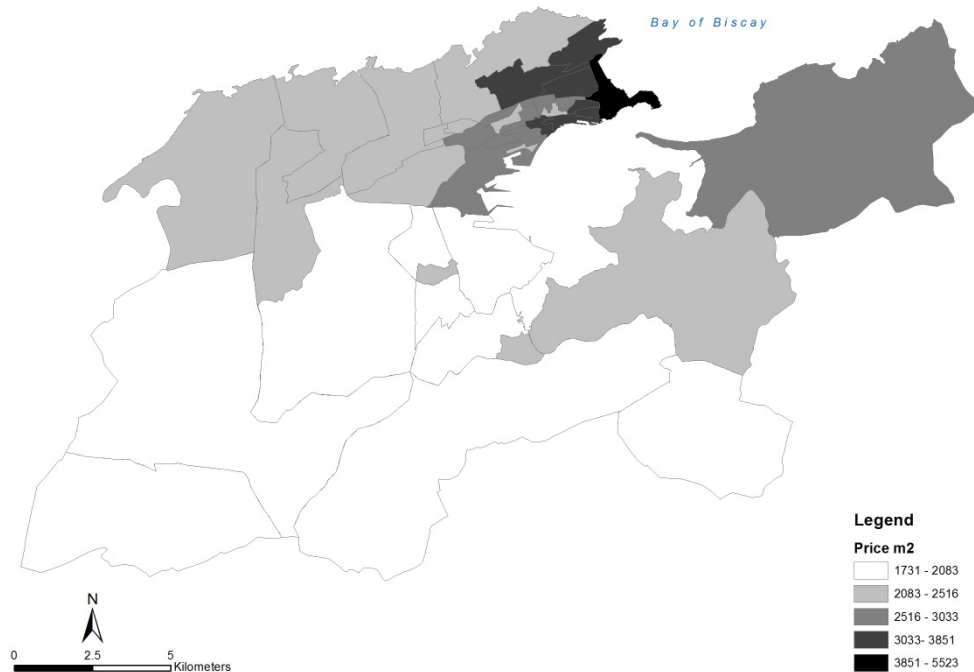
### 4.1 Available data and study areas

Hedonic regression models have been estimated using two databases from the cities and areas of influence around Rome and Santander. Aggregated data was available for both areas from a total of 211 zones in the case of Rome and 42 zones in the case of Santander. The study areas have clearly differentiated characteristics. Rome and its surrounding area contain almost two and a half million inhabitants and a million jobs. The urban area has a primary centre formed of traditional historic neighbourhoods (Municipality I) and two secondary centres in the business area EUR and the port of Ostia. Together these three locations contain 20% of the jobs in the overall study area. The city also has serious problems with mobility and traffic congestion. Santander, on the other hand, is a medium sized city with a surrounding area made up of 8 municipalities with a total of 260,000 inhabitants and a little over 100,000 jobs. The area has a clear monocentric nature with the main nucleus of Santander concentrating 20% of the jobs. The city currently has moderate problems associated with congestion mainly at the main access points to the city centre.

Figure 1 shows the average residential property prices (in euros per square meter) in each of the areas. In the case of Rome, the prices follow a clearly concentric pattern from the historic centre outwards towards the periphery. In Santander the prices follow a north – south gradient towards inland from the coast and from east to west in the case of the urban nucleus.







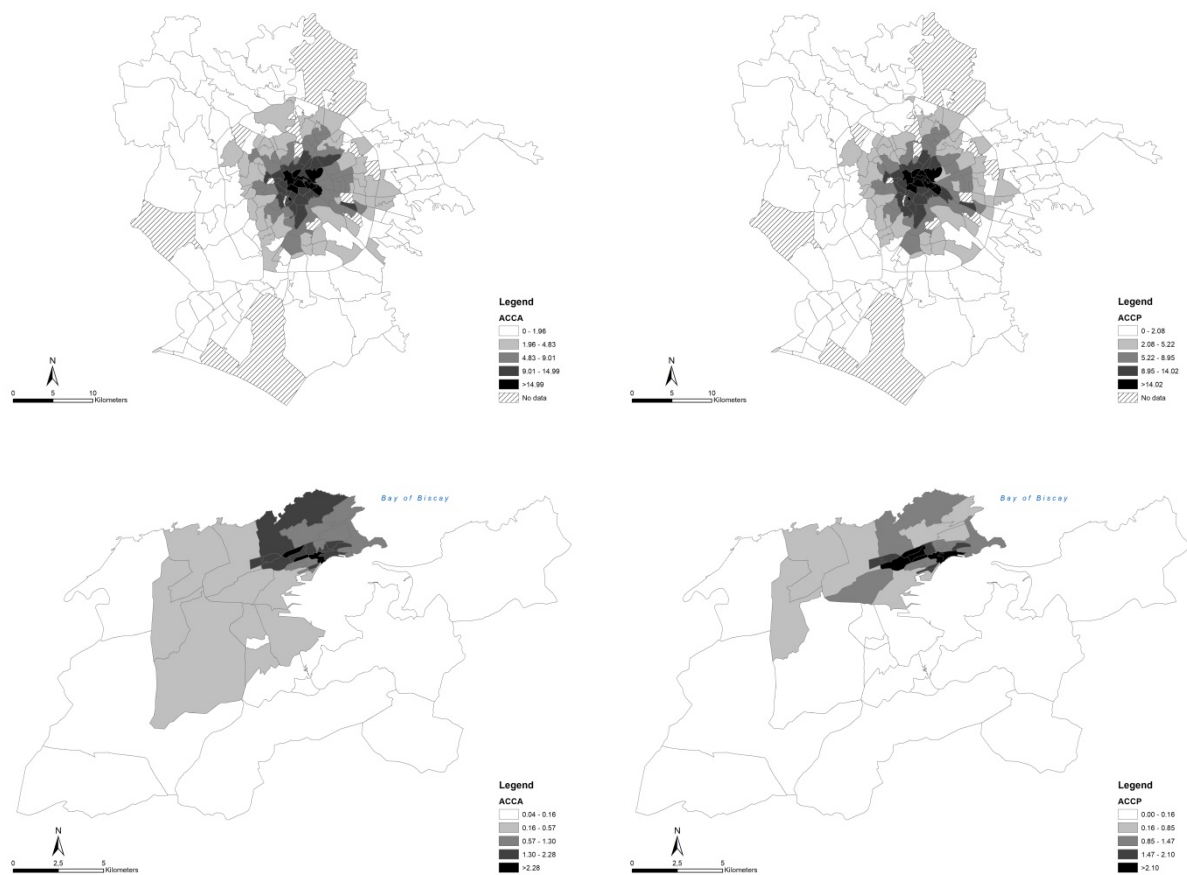
**Figure 1- Real estate prices (euros/m<sup>2</sup>) in the cities of Rome (above) and Santander (below)**

The variables contained in the database for both cities can be consulted in Table 2 and Table 3. The prices per square meter (variable P) were clearly higher in Rome than in Santander with an average of 3,440 euros/m<sup>2</sup> versus 2,661 euros/m<sup>2</sup>. Furthermore, the price range in Rome was also wider with a maximum of 8,250 euros/m<sup>2</sup> compared with 5,523 euros/m<sup>2</sup> in Santander. In both cases, the prices correspond to offer prices obtained from real estate sites collected in 2010. These offer prices are usually 1 to 10 percent higher than selling prices with which they are highly correlated (Chasco and Le Gallo, 2015; Hometrack, 2017).

The variables POP, HOUSE, SQM, EMP, AGE and APART measure the population, the number of dwellings, the residential square meters, the number of jobs, the average age of the buildings and the proportion of apartments in each zone respectively. The first four variables have been considered as densities taking into account the surface of each of the zones, which are usually smaller in city centres and bigger in the outskirts. The existing jobs in each area have been collected, in the case of Rome, from the Population Census, while in Santander, given that the Spanish Population Census does not collect this information, from a regional directory of companies. The variable AGE considers the average age of the buildings in four classes: 1) before 1900, 2) between 1900 and 1945, 3) between 1946 and 1977 and 4) after 1977. APART is the proportion of apartments with respect to the total dwellings of the zone, i.e. apartments plus single houses.

The indicators of active and passive accessibility (ACCA and ACCP) using public

transport are respectively specified using the jobs and population present in the zones and in both cases the indicators had a lower value in Santander because of the lower number of jobs and population present. Indicators (6) and (7) were estimated using as dependent variables the trips produced by each zone in the case of active accessibility and the trips attracted by each zone in the case of passive accessibility. In this way, the production and attraction of trips obtained from the trip matrices of the study areas, estimated using a double constrained gravity model, act as a proxy for the active and passive accessibility of the zones. The spatial distribution of accessibility indicators are shown in Figure 2. It can be seen that there is a strong correlation between active and passive accessibility and how the higher values of accessibility are reached in the main centres in both, Rome and Santander. The parameters calibrated for the indicators can be seen in Table 1.



**Figure 2- Active (left) and Passive (right) accessibility in the cities of Rome (above) and Santander (below)**

The journey time by public transport does not only include the time on board the vehicle, but also the waiting and transfer time, if required, for each zone. The public transport considered for Rome in this research was the bus and metro systems whereas in Santander it was the urban and inter-urban rail and bus services. Public transport was

chosen as an indicator of cost rather than using the travel times taken by private transport due to its greater availability for all types of users and given that we wanted to focus on value capture policies based on the capitalisation of the value generated by public transport in real estate prices.

Parameter	Rome	Santander
$\alpha_1$	0.05	0.02
$\alpha_2$	-0.1	-0.1
$\alpha_3$	0.85	0.02
$\alpha_4$	-1.22	-0.1

**Table 1. Parameters calibrated for the accessibility indicators**

The journey time to the city centre (CBD) in minutes using public transport was specified in the case of Rome considering only the area containing the historic neighbourhoods as the centre, whereas in Santander it was located in the main city centre. The indicator was, once again, higher in the case of Rome, with an average value greater than 67 minutes compared to 49 minutes for Santander. The public transport spatial coverage variable (TRANS) was higher in the case of Santander with an average of 73% of the area being covered, although it must be remembered that the Santander study area is considerably smaller than the area of Rome (299 Km<sup>2</sup> in the case of Santander compared to 1,326 Km<sup>2</sup> in the case of Rome). In both cases coverage was measured using a 400 meter radius around the public transport stops considering a maximum of 6 minutes access time for a pedestrian speed of 4km/h.

The rest of the control variables are dummies taking a value of 1 if the zone is part of the urban centre (CEN), is part of a commercial area (COM), a green zone (GREEN) or if it belongs to a prestigious area (PG), the latter being a variable of a particularly subjective nature which needs to be determined by the analyst through a qualitative analysis of the study area. Note that in the case of the dummy variables, the proportion of the zones showing each of the characteristics is very similar in both urban areas, although the zones belonging to the centre and, above all, to a commercial area are more numerous in Rome than in Santander.

Variable	Description	Units	Mean	Standard deviation	Minimum	Maximum	Expected sign
P	Property price per square meter	Euros	3,440.40	1,208.81	1,950	8,250	
POP	Population/Km <sup>2</sup>	-	11,790.5	9,738.86	274.27	49,693.0	+/-

HOUSE	Dwellings/Km <sup>2</sup>	-	2,904.43	3,279.25	14.38	18,179.6	+/-
SMQ	Built square meters/square meters	%	24.18	25.74	0.13	123.3	+/-
EMP	Jobs/Km <sup>2</sup>	-	5,584.83	6,160.2	158.75	36,794.0	+/-
AGE	Average age of the buildings	1-4	3.07	0.74	1	4	+/-
APART	Proportion of apartments	0-1	0.98	0.07	0.25	1	+/-
ACCA	Active accessibility using public transport	-	4.34	4.74	0	23.48	+
ACCP	Passive accessibility using public transport	-	4.44	4.68	0	21.43	+
CBD	Travel time to city centre using public transport	Minutes	67.04	28.69	13.26	166.10	-
TRANS	Public transport coverage in the zone	0-1	0.19	0.26	0	1	+
CEN	Belonging to the central zone	1/0	0.07	0.25	0	1	+
COM	Belonging to a commercial zone	1/0	0.06	0.20	0	1	+/-
GREEN	Presence of green areas	1/0	0.14	0.33	0	1	+
PG	Zone of special prestige	1/0	0.12	0.31	0	1	+

**Table 2 – Descriptive statistics of the variables contained in the database of Rome (N=211 zones)**

Variable	Description	Units	Mean	Standard deviation	Minimum	Maximum	Expected sign
P	Property price per square meter	Euros	2,661.31	716.31	1,731	5,523	
POP	Population/Km <sup>2</sup>	-	13,748.1	17,612.59	117.46	65,555.55	+/-
HOUSE	Dwellings/Km <sup>2</sup>	-	4,992.24	6,663.85	25.77	24,625.6	+/-
SMQ	Built square meters/square meters	%	40.45	54.21	0.29	209.1	+/-
EMP	Jobs/Km <sup>2</sup>	-	4,559.53	8,463.14	12.02	35,060.6	+/-
AGE	Average age of the buildings	1-4	3.24	0.43	3	4	+/-
APART	Proportion of apartments	0-1	0.75	0.28	0.03	1	+/-
ACCA	Active accessibility using public transport	-	0.97	0.92	0.04	3.56	+
ACCP	Passive accessibility using public transport	-	0.99	0.95	0	3.36	+

CBD	Journey time to the city centre by public transport	Minutes	49.31	20.69	13.45	82.24	-
TRANS	Public transport coverage in the zone	0-1	0.73	0.33	0.08	1	+
CEN	Belonging to the central zone	1/0	0.05	0.22	0	1	+
COM	Belonging to a commercial zone	1/0	0.12	0.33	0	1	+/-
GREEN	Presence of a green zone	1/0	0.14	0.35	0	1	+
PG	Zone of special prestige	1/0	0.10	0.30	0	1	+

**Table 3 – Descriptive statistics of the variables contained in the database of Santander (N=42 zones)**

## 4.2 Results of the models

The parameters estimated for the models of Rome and Santander are summarised in Table 4 and Table 5. The R-1, R-2, S-1 and S-2 models correspond to the hedonic regressions of Rome (R) and Santander (S), respectively, whereas the R-3, R-4, S-3 and S-4 models are autoregressive spatial models considering spatial dependence in the property prices. The variable price/m<sup>2</sup> in all cases has been specified in a logarithmic form. This functional form has various advantages such as the reduction of possible heteroscedasticity in the residuals of the models (Malpezzi, 2008). The logarithmic specification of the dependent variable allows us to interpret the estimated parameters as semi-elasticities, in other words, the percentage change of the dependent variable as a unit change in the independent variable whilst maintaining the rest of the variables as constants. In the case of the dummy variables, the effect of the parameter should be estimated by using the expression:  $[\exp(\beta_n) - 1] * 100$ , where  $\beta_n$  is the parameter of the dummy variable being considered (Halvorsen and Palmquist, 1980).

The specification of the models has excluded variables showing high correlation between each other and which could, therefore, generate problems of collinearity in the models. Firstly, the variables POP, HOUSE and SQM showed a Pearson correlation coefficient greater than 0.8 which meant that only the POP variable was chosen for inclusion in the models. Furthermore, the functions were specified either with the gravity type accessibility indicator, or with the indicator of travel time to the city centre to detect which one had a better fit to the data. Both indicators were not used simultaneously as they had a high correlation (greater than 0.6). Furthermore, ACCA and ACCP were also strongly correlated in both study areas so only active accessibility was chosen to evaluate the influence access to opportunities, in this case measured using the number of jobs, had on real estate values. The strong correlation between ACCA and ACCP was detected from their correlation coefficient greater than 0.8 and the Variance Inflation Factors greater than 50 if both variables were specified in the model at the same time.

The R-1 and S-1 models were finally specified with a total of 10 independent variables adopting the following functional form:

$$\ln(P_i) = \beta_1 + \beta_2 POP_i + \beta_3 EMP_i + \beta_4 AGE_i + \beta_5 APART_i + \beta_6 ACCA_i + \beta_7 TRANS_i + \beta_8 CEN_i + \beta_9 COM_i + \beta_{10} GREEN_i + \beta_{11} PG_i + \varepsilon_i \quad (8)$$

The fit of the R-1 model was 74% of the explained variance. The population density of the zone (POP variable) had a small negative effect but it was clearly not significant. The jobs present in the zone (EMP) and the proportion of apartments (APART) were not significant either. The average age of the buildings (AGE) presented a negative sign which indicates the greater value of the historical buildings in the city of Rome. Among the dummy variables that considered different zonal characteristics, only the presence of special prestige in the area was significant. The presence of green zones had a negative sign probably due to a higher availability of these facilities in peripheral areas, while prestige was a clearly positive and important factor with an increase of 37% in the average property prices of the zone. The gravity accessibility variable was significant with an increase of 3.5% in property values per additional unit in the indicator. The coverage of public transport in the zone (TRANS) presented a negative sign and it was not significant.

This model was also specified with the travel time on public transport to the historic urban centre (CBD) instead of the gravity indicator (R-2 model). In this case, the goodness of fit of the model increased to 0.78 in the  $R^2$  indicator. The number of jobs variable was positive and significant, while the coverage of public transport stops showed in this case a parameter clearly not significantly different from zero. Among the dummy variables which considered environmental characteristics, the presence of green zones changed sign and continued being a parameter non-significantly different from zero, whereas CEN presented a clearly positive and significant parameter. It can generally be said, therefore, that the relative accessibility indicator slightly improved the fit and the coherence of the parameters compared with the gravity accessibility indicator, although both showed that accessibility to opportunities has a positive impact on real estate prices.

The residuals of the R-1 and R-2 models were tested using the Moran I index to check for a significant level of spatial autocorrelation. This autocorrelation was found to be positive and strong in both models. To determine which spatial model could be more suitable for capturing this correlation, a robust Lagrange Multipliers (LM) test was applied and found that the LM – lag test for the SAR model was clearly significant and showed very superior values to the LM – error test for the SEM model.

Given these results, the R-3 and R-4 models were specified with the same variables as

the R-1 and R-2 models but by also considering the presence of spatial dependence in the price of neighbouring properties (SAR model). The chosen neighbourhood matrix was composed of the 5 closest zones. This alternative was chosen over other configurations because it guarantees a homogenous distribution of neighbourhood between observations, which is recommended as the calculation of the asymptotic properties of the estimators and the statistical tests are based on the hypothesis of the existence of regularity (Anselin, 2002).

Variable	R-1	R-2	R-3	R-4
(Intercept)	7.919 (.000)	8.433 (.000)	2.473 (.000)	3.143 (.000)
POP	-0.000 (.828)	0.000 (.949)	-0.000 (.786)	-0.000 (.867)
EMP	0.000 (.481)	0.000 (.000)	0.000 (.771)	0.000 (.051)
AGE	-0.051 (.030)	-0.049 (.022)	-0.012 (.483)	-0.013 (.410)
APART	0.159 (.345)	0.108 (.484)	0.104 (.377)	0.088 (.452)
ACCA	0.035 (.000)	-	0.012 (.002)	-
CBD	-	-0.005 (.000)	-	-0.002 (.000)
TRANS	-0.111 (.068)	-0.088 (.105)	-0.031 (.468)	-0.030 (.467)
CEN	0.073 (.133)	0.123 (.006)	0.049 (.149)	0.072 (.032)
COM	-0.026 (.638)	-0.071 (.155)	-0.015 (.695)	-0.035 (.358)
GREEN	-0.033 (.368)	0.063 (.071)	-0.037 (.144)	0.003 (.910)
PG	0.317 (.000)	0.262 (.000)	0.172 (.000)	0.161 (.000)
$\rho$	-	-	0.677 (.000)	0.620 (.000)
R <sup>2</sup>	0.74	0.78	-	-
Adjusted R <sup>2</sup>	0.73	0.77	-	-
F Test	58.37 (.000)	72.71 (.000)	-	-
Log-likelihood	97.60	115.33	156.61	161.91
AIC	-173.21	-208.66	-287.22	-297.83
Moran I	0.36 (.000)	0.31 (.000)	0.01 (.719)	0.02 (.555)
Robust LM – Lag	40.77 (.000)	34.10 (.000)	-	-
Robust LM – Error	2.42 (.120)	2.06 (.151)	-	-
LR test	-	-	118.01 (.000)	93.16 (.000)

**Table 4 – Hedonic models estimated for the area of Rome (in brackets the p – value with the statistical significance of the parameters)**

The parameters estimated in the R-3 and R-4 models were generally lower in magnitude than those estimated in R-1 and R-2, although the overall impacts of each variable need to be considered, as will be described in the following section. Nevertheless, both the accessibility indicator and the travel time to the centre indicator showed significant parameters with estimations of 1.2% increases in average property prices in the zones per additional unit of accessibility or a reduction of 0.2% per additional minute in the travel time to the urban centre by public transport. The presence of greater spatial coverage of public transport continued to be a variable without significance and the average age of the buildings ceased to be significant.

The likelihood ratio (LR) test between the SAR models and the conventional hedonic models were clearly significant showing that R-3 and R-4 had a clearly better fit to the data. In both cases the estimated  $\rho$  parameters were clearly significant and the Moran I statistic performed on the residuals of the spatial models showed how autocorrelation in the residuals stopped being present.

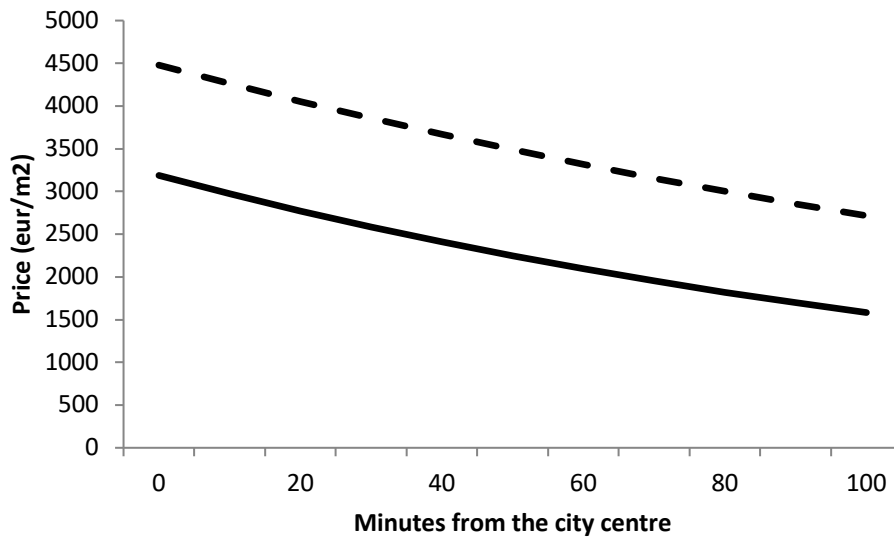
Variable	S-1	S-2	S-3	S-4
(Intercept)	7.502 (.000)	8.027 (.000)	3.192 (.002)	4.111 (.000)
POP	-0.000 (.091)	-0.000 (.008)	-0.000 (.002)	-0.000 (.000)
EMP	0.000 (.227)	0.000 (.043)	0.000 (.019)	0.000 (.001)
AGE	-0.025 (.666)	-0.015 (.769)	0.009 (.828)	0.013 (.728)
APART	0.299 (.068)	0.260 (.076)	0.203 (.081)	0.185 (.083)
ACCA	0.054 (.189)	-	0.031 (.293)	-
CBD	-	-0.007 (.003)	-	-0.005 (.002)
TRANS	0.213 (.140)	0.052 (.707)	0.188 (.063)	0.064 (.527)
CEN	-0.410 (.090)	-0.601 (.010)	-0.437 (.010)	-0.588 (.000)
COM	-0.050 (.586)	0.023 (.782)	0.043 (.529)	0.086 (.176)
GREEN	0.261 (.003)	0.238 (.003)	0.139 (.034)	0.137 (.022)
PG	0.267 (.041)	0.200 (.090)	0.194 (.034)	0.151 (.076)
$\rho$	-	-	0.552 (.000)	0.484 (.000)
R <sup>2</sup>	0.77	0.81	-	-
Adjusted R <sup>2</sup>	0.69	0.75	-	-
F Test	10.16 (.000)	13.51 (.000)	-	-



Log-likelihood	31.59	36.34	38.08	41.98
AIC	-41.18	-50.70	-50.17	-57.97
Moran I	0.12 (.004)	0.08 (.024)	0.04 (.408)	0.02 (.559)
LM – Lag robust	12.48 (.000)	11.73 (.000)	-	-
LM – Error robust	1.50 (.220)	1.40 (.237)	-	-
LR test	-	-	13.03 (.000)	11.34 (.000)

**Table 5– Hedonic models estimated for the area of Santander (in brackets the p – value with the statistical significance of the parameters)**

The models estimated for the case of Santander were specified with the same variables used for Rome and therefore with the same neighbourhood matrix in the case of the specifications considering spatial autocorrelation, S-3 and S-4. In the models for Santander, the parameters estimated in the ACCA variable, were a little higher than those estimated for the Rome models, even though they were not clearly significant in this case. Nevertheless, according to the t test, the parameters estimated for the distance to the CBD were significantly different from zero with a value between 0.7% and 0.5% reduction in the average zonal price of properties per additional minute of travel time to the city centre. The TRANS variable, contrary to the results found in Rome, was more clearly significant, especially in the S-1 and S-3 models, although in the cases specified with the CBD variable the significance was reduced. The average age of the buildings was a clearly not significant variable in Santander whereas the proportion of apartments had a positive sign and was nearly significant (probably due to the lower prices of the rural houses inside the study area). Other environmental variables which proved to be relevant were the presence of green zones, with positive impacts of around 15% to 30% increase in average property prices, and also PG with an increase between 16% to 31 depending on the model.



**Figure 3– Partial effect of the CBD variable in the R-2 (dotted line) and S-2 (solid line) models**

The goodness of fit for the Santander models was similar, according to  $R^2$ , as it was for Rome. However, the presence of spatial autocorrelation in the residuals of the models was weaker and the LR tests provided lower values, which imply that the SAR models improved less the fit when compared with the simple hedonic regression models.

Figure 3 shows a comparison between the partial effects of the CBD variable for the R-2 and S-2 models. It can be seen how, given the higher effect of the variable in the case of Santander, the fall in property prices does not converge for high travel times with the fall in prices of Rome.

### 4.3 Overall impacts in the spatial autoregressive models

In order to correctly interpret the parameters of the SAR models, the overall impacts of the variables are calculated (see Table 6). These overall impacts take into account both the direct effect of each independent variable on the dependent variable and the indirect effect of spatial diffusion provoked on the neighbouring observations and from these on the local observation (LeSage and Pace, 2010b).

These total effects are calculated from the partial derivatives matrix for each independent variable. This matrix presents in its main diagonal the parameters corresponding to the direct impacts from each observation in the dependent variable, while the parameters outside the main diagonal represent the indirect impacts in neighbouring observations due to the effects of spatial dependency.

Taking into account these direct and indirect effects, it can be seen how the SAR models allow us to infer that the overall effect caused by changes made to certain variables, including ACCA and CBD, is greater than that found by directly estimating the hedonic model. Considering these overall impacts, an additional minute of travel time to the city centre using public transport in Rome implied a drop of 0.6% in property prices and of 1% in Santander, in other words, 0.5% more than when indirect impacts are not considered. In the case of Rome, the gravity accessibility indicator was also significant with an increase of 3.8% in average property values per additional unit, a value which is clearly greater than the 1.2% estimated without considering the indirect impacts.

Variable	R-3	R-4	S-3	S-4
POP	-0.000 (.860)	-0.000 (.835)	-0.000 (.059)	-0.000 (.017)
EMP	0.000 (.791)	0.000 (.027)	0.000 (.085)	0.000 (.034)
AGE	-0.036 (.495)	-0.035 (.417)	0.020 (.835)	0.026 (.814)
APART	0.321 (.463)	0.231 (.418)	0.452 (.152)	0.358 (.084)
ACCA	0.038 (.001)	-	0.069 (.327)	-
CBD	-	-0.006 (.000)	-	-0.010 (.021)
TRANS	-0.096 (.397)	-0.078 (.539)	0.419 (.124)	0.125 (.681)
CEN	0.151 (.175)	0.190 (.021)	-0.974 (.085)	-1.140 (.022)
COM	-0.046 (.880)	-0.091 (.451)	0.096 (.611)	0.168 (.276)
GREEN	-0.115 (.159)	0.008 (.992)	0.310 (.038)	0.266 (.024)
PG	0.533 (.000)	0.424 (.000)	0.434 (.176)	0.293 (.123)

**Table 6– Parameters estimated with the overall impacts of the SAR models**

## 5. DISCUSSION AND CONCLUSIONS

This research has estimated hedonic regression models and spatial hedonic regression models to determine the presence and magnitude of the impacts accessibility to opportunities using public transport has on property prices in the cities of Rome and Santander.

The results confirm that accessibility to opportunities was a significant factor on property price increases in both study areas. However, whereas in the case of Rome both the gravity indicator and the relative accessibility to the urban centre captured this effect, the travel time to the CBD was the only significant variable in the case of Santander. This fact is almost certainly due to the monocentric nature of the Santander

urban area compared to that of Rome where the sphere of influence contains two important sub-centres apart from the traditional historical central area. The coverage of public transport stops did not prove to be a relevant variable in the case of Rome, and it was not shown to be much more important in Santander, given that when it was specified along with the CBD variable it was clearly not significant. All of which indicates the relevance of using an accessibility indicator fitted to the study area in order to effectively capture the contribution this factor makes to property prices. Furthermore, it is convenient to consider not only accessibility to transport services, which has been normal practise in most research, but also accessibility to end opportunities, as that is what motivates the trips made by users.

In terms of the effects of the possible presence of spatial dependence in the real estate market, in Rome the autocorrelation of the hedonic model residuals was clearly significant and strong. However, it was weaker in the area of Santander and the SAR models did not improve too much the goodness of fit to the data over that of the conventional hedonic models. The better fit of the spatial models with the Rome data appears to be due to the presence of diffusion effects and the inter-zonal influence on property prices, a phenomenon which occurred less markedly in Santander, possibly because the effects of diffusion were neutralised by zonal aggregation. Finally, the calculation of overall impact by the SAR models allowed us to test whether the effects of accessibility were greater than those estimated by the non-spatial hedonic regression models. This effect was greater in the case of Santander than in Rome if travel time to the CBD is considered as the prices showed a drop of 0.6% per minute of additional travel time to the centre of Rome compared with 1% in the case of Santander. This result agrees with expectations, given that the size of the Santander core is much smaller than the center of Rome. In addition, these estimates are similar to those provided by Marmolejo and González (2009) for the case of Barcelona.

The way accessibility is seen to be capitalised in real estate prices supports the possibility of using value capture policies in the study areas to provide increased finance for improving public transport. Such policies have already been successfully introduced in a variety of urban areas, providing further support for their introduction in new cities.

## **ACKNOWLEDGEMENTS**

This research was made possible through the Project PARK-INFO (TRA2013-48116-R) financed by the Spanish Ministry of Economy and Competitiveness.

## **REFERENCES**

Anselin, L. (1988) *Spatial econometrics : methods and models*. Kluwer Academic

- Publishers, Dordrecht; Boston.
- Anselin, L. (2002) Under the hood : Issues in the specification and interpretation of spatial regression models. *Agricultural Economics* 27, 247-267.
- Anselin, L., Rey, S.J. (2014) *Modern spatial econometrics in practice: A guide to GeoDa, GeoDaSpace and PySAL*.
- Armstrong, R., Rodriguez, D. (2006) An Evaluation of the Accessibility Benefits of Commuter Rail in Eastern Massachusetts using Spatial Hedonic Price Functions. *Transportation* 33, 21-43.
- Arnott, R.J., Stiglitz, J.E. (1979) Aggregate land rents, expenditure on public goods, and optimal city size. *The Quarterly Journal of Economics*, 471-500.
- Bitter, C., Mulligan, G., Dall'erba, S. (2007) Incorporating spatial variation in housing attribute prices: a comparison of geographically weighted regression and the spatial expansion method. *J Geograph Syst* 9, 7-27.
- Bowes, D.R., Ihlanfeldt, K.R. (2001) Identifying the Impacts of Rail Transit Stations on Residential Property Values. *Journal of Urban Economics* 50, 1-25.
- Boyle, M., Kiel, K. (2001) A Survey of House Price Hedonic Studies of the Impact of Environmental Externalities. *Journal of Real Estate Literature* 9, 117-144.
- Cascetta, E. (2009) *Transportation systems analysis : models and applications*, 2nd ed. Springer, New York.
- Cervero, R., Kang, C.D. (2011) Bus rapid transit impacts on land uses and land values in Seoul, Korea. *Transport Policy* 18, 102-116.
- Colombo, E., Stanca, L. (2014) Measuring the monetary value of social relations: A hedonic approach. *Journal of Behavioral and Experimental Economics* 50, 77-87.
- Coppola, P., Nuzzolo, A. (2011) Changing accessibility, dwelling price and the spatial distribution of socio-economic activities. *Research in Transportation Economics* 31, 63-71.
- Court, A.T. (1939) Hedonic Price Indexes With Automobile Examples. *The dynamics of automobile demand*. General Motors, New York.
- Chasco, C., Le Gallo, J. (2015) Heterogeneity in Perceptions of Noise and Air Pollution: A Spatial Quantile Approach on the City of Madrid. *Spatial Economic Analysis* 10, 317-343.
- Chen, H., Rufolo, A., Dueker, K. (1998) Measuring the impact of light rail systems on single-family home values: a hedonic approach with geographic information system application. *Transportation Research Record: Journal of the Transportation Research Board*, 38-43.
- Debrezion, G., Pels, E., Rietveld, P. (2007) The Impact of Railway Stations on Residential and Commercial Property Value: A Meta-analysis. *The Journal of Real Estate Finance and Economics* 35, 161-180.
- Dubé, J., Thériault, M., Des Rosiers, F. (2013) Commuter rail accessibility and house values: The case of the Montreal South Shore, Canada, 1992–2009. *Transportation Research Part A: Policy and Practice* 54, 49-66.
- Efthymiou, D., Antoniou, C. (2013) How do transport infrastructure and policies affect house prices and rents? Evidence from Athens, Greece. *Transportation Research Part A: Policy and Practice* 52, 1-22.
- Fogarty, N.A., America, R. (2008) *Capturing the value of transit*. Center for Transit-Oriented Development.
- Fotheringham, A.S., Brunson, C., Charlton, M. (2002) *Geographically weighted regression : the analysis of spatially varying relationships*. Wiley, Chichester, England ; Hoboken, NJ, USA.
- François, D.R., Antonio, L., Marius, T., Marcel, B. (1996) Shopping centres and house

- values: an empirical investigation. *Journal of Property Valuation and Investment* 14, 41-62.
- Halvorsen, R., Palmquist, R. (1980) The Interpretation of Dummy Variables in Semilogarithmic Equations. *American Economic Review* 70, 474-475.
- Handy, S.L., Niemeier, D.A. (1997) Measuring accessibility: An exploration of issues and alternatives. *Environment and Planning A* 29, 1175-1194.
- Hometrack (2017) UK Cities House Price Index, December 2017. UK House Price Statistics.
- Ibeas, A., Cordera, R., dell'Olio, L., Coppola, P., Dominguez, A. (2012) Modelling transport and real-estate values interactions in urban systems. *Journal of Transport Geography* 24, 370-382.
- Jim, C.Y., Chen, W.Y. (2010) External effects of neighbourhood parks and landscape elements on high-rise residential value. *Land Use Policy* 27, 662-670.
- Jojoa, J.P., Marmolejo, C. (2013) Evaluación del impacto de la estratificación socioeconómica en la formación del valor del suelo en Bogotá D.C. 2013.
- Koenig, J.G. (1980) Indicators of urban accessibility: Theory and application. *Transportation* 9, 145-172.
- LeSage, J.P., Pace, R.K. (2009) *Introduction to spatial econometrics*. CRC Press, Boca Raton.
- Lesage, J.P., Pace, R.K. (2010a) The Biggest Myth in Spatial Econometrics. Available at SSRN: <http://ssrn.com/abstract=1725503> or <http://dx.doi.org/10.2139/ssrn.1725503>.
- LeSage, J.P., Pace, R.K. (2010b) Spatial Econometric Models. *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications*. Springer, Berlin, pp. 355-376.
- Malpezzi, S. (2008) Hedonic Pricing Models: A Selective and Applied Review. *Housing Economics and Public Policy* ed Tony O'Sullivan, K.G., pp. 67-89.
- Marmolejo, C., González, C. (2009) Does noise have a stationary impact on residential values? *Journal of European Real Estate Research* 2, 259-279.
- Martínez, F.J. (2000) Towards a Land-use and transport interaction framework. *Handbook of Transport Modelling* eds Hensher, D.A., Button, K.J. Elsevier Science, pp. 145-164.
- Martínez, L., Viegas, J. (2009) Effects of Transportation Accessibility on Residential Property Values. *Transportation Research Record: Journal of the Transportation Research Board* 2115, 127-137.
- Mohammad, S.I., Graham, D.J., Melo, P.C., Anderson, R.J. (2013) A meta-analysis of the impact of rail projects on land and property values. *Transportation Research Part A: Policy and Practice* 50, 158-170.
- Munoz-Raskin, R. (2010) Walking accessibility to bus rapid transit: Does it affect property values? The case of Bogotá, Colombia. *Transport Policy* 17, 72-84.
- Panduro, T.E., Veie, K.L. (2013) Classification and valuation of urban green spaces—A hedonic house price valuation. *Landscape and Urban Planning* 120, 119-128.
- Rosen, S. (1974) Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy* 82, 34-55.
- Senior, M.L. (2009) Impacts on travel behaviour of Greater Manchester's light rail investment (Metrolink Phase 1): evidence from household surveys and Census data. *Journal of Transport Geography* 17, 187-197.
- Wang, Y., Kockelman, K.M., Wang, X. (2013) Understanding spatial filtering for analysis of land use-transport data. *Journal of Transport Geography* 31, 123-131.