Does urban centrality influence residential prices? An analysis for the Barcelona Metropolitan Area

¿Incide la centralidad urbana en la formación de los precios residenciales? Un análisis para la Barcelona Metropolitana

Carlos Marmolejo-Duarte (Main and Correspondant Author)

Departamento de Tecnología en la Arquitectura, Centro de Política de Suelo, Escuela Técnica Superior de Arquitectura de Barcelona, Universidad Politécnica de Cataluña
Diagonal, 649, 4a, Arq Legal, 08028, Barcelona, Spain
Carlos.marmolejo@upc.edu

Manuscript Code: 638
Date of Acceptance/Reception: 27.03.2017/21.05.2015.
DOI: 10.7764/RDLC.16.1.57

Abstract
The bid rent theory (BRT), originally conceived for a monocentric city suggests a trade-off between land value and transport costs. Thus, in most of the practical applications, the simply distance/time/cost to the CBD is used as a proxy of accessibility. Nonetheless, in contemporary metropolises, employment and services do not cluster in one CBD but in many centers, furthermore the centrality quality does not follow a smooth gradient as distance to centers increases. Consequently, taking the distance/time/cost to centers in the context or hedonic models is problematic for collinearity issues and too simplistic since it directly assumes a smoothed gradient function. In this paper we test in Barcelona Metropolitan Area, a very well recognized polycentric city, whether some continuous indicators of centrality are key determinants of housing prices. Using listing prices, a hedonic model is built, and the asking price is regressed over two continuous indicators of centrality, one of them calculated departing of the spatial-temporal behavior of people, which itself is a novelty in this kind of studies. The results suggest that continuous centrality indicators do exert a moderate influence on housing prices after controlling for other structural and locative attributes. Nevertheless, the main determinants of prices are related to the socioeconomic stratification not accessibility as suggested by BRT. Energy class appears also as a factor influencing dwellings’ price.

Key words: Hedonic price models, real estate valuation, polycentrism, time-geography, Barcelona.

Introduction
The “standard urban model” based in the bid rent theory, with roots in the von Thünen’s (1826) and Hurd’s (1924) works, as it was depicted by Alonso (1965), explains that both land rent and urban density declines as the distance to center increases. In such model, it is assumed that most of the employment is concentrated in the center, which foster a daily commuting between peripheral residential zones and it. By having the employment clustered in the center it also is assumed that services follow such spatial trend. All in all, result in increasing transport cost and time (Haig, 1926) as distance to center increases, and consequently a trade off with rent allocated to residential land. It is the existence of the land rent gradient that underlies the formation of density in a competitive market scenario. If it is considered that developers invest capital in land and building when developing a site, and constant returns per unit of land are relaxed (i.e. once substitution between land and building costs is allowed), real estate developers economize on the use of expensive land in more central locations where prices peak. In optimizing developments, they add more building capital per unit of capital invested on land (i.e. they build multi-story structures instead of low rise ones), despite the fact that
high rise building is costlier per sq. meter than low rise ones. Nonetheless, the capital saved in land outbids the over cost invested in the construction. Consequently, it emerges a density gradient following that of land prices. The parallelism between rent and density gradients depends upon the elasticity of substitution between land and capital (Kau & Lee, 1976). Mills and Hamilton (1984) demonstrated, starting from the monocentric city model, that under certain constraints, such as Cobb-Douglas’s production function for housing, users with identical tastes and income and unit price elasticity of demand for housing, density gradients adopt a negative exponential function. According to the above state theory, built densities are proxies for centrality, and theoretically, if land rent does not only reflects accessibility but all the remaining local amenities, built densities also proxies for them and not only for centrality.

Description of the problem

Nonetheless, monocentric metropolises do not more exist and probably have never existed as they were stylized in theoretical models (Anas et al., 1998; Batty, 2001). Furthermore, employment is not the only driver of urban mobility and recent studies have reported a decrease of its relative importance in relation to non-obligated mobility (Cerda & Marmolejo, 2010). On the other hand, taking the distance to subcenters and/or CBD is problematic itself for a number of reasons:

1) Firstly, it is necessary to previously identify subcenters and validate them which is itself problematic (McDonald, 1987)
2) Secondly, it is not possible to identify clear limits of centralities, since centrality quality is an attribute that has a spatial fading (centrality gradient) departing from centers, even in very well planned cities where exclusive zoning is implemented
3) Thirdly, the centrality gradient is not regular enough to be substituted by the simply distance to centers; thought it is always possible to use more sophisticated locally adjusted regressions models, like those used by Marmolejo & González (2009)
4) Fourthly, in polycentric metropolitan areas, where multiple centers give structure to quotidian flows, taking as many distances as subcenters there are results into multicollinearity when hedonic prices are calibrated (Aguirre & Marmolejo, 2010). Conversely, taking only the distance to the nearest center neglect the fact that sites may receive the multiple influence of neighboring subcenters not only the nearest (García-López & Muñiz, 2010).

For the above stated issues, in this paper we test whether 2 continuous indicators of centrality, one based on built density and other on the spatial-temporal behavior of population (depicting all urban activities not only labor ones), are able to explain the spatial trends of housing prices. In doing so, a hedonic price model (Rosen, 1974) is built departing from listing prices for multifamily dwellings in Metropolitan Barcelona, taking as explanatory variables both structural attributes and locative ones. In this latter continuous indicators of centrality are included, of them that departing from spatial-temporal population trends is a novelty by itself in this kind of studies.

The remaining of the paper is organized as follows. First, a brief literature revision is offered; followed by the presentation of data, case study and methodology used; the results and discussion is presented next; in the concluding epigraph, the main findings are highlighted.

Brief literature review

For analytical purposes real estate values, can be easily split in land rent and the value related to the building (quality, area, design, facilities, etc.). In urban economics land rent is basically associated to accessibility to centers where services and employment is supposed to be concentrated (O’Sullivan, 2007).

In Barcelona, the pioneering work of Josep Roca (1988) uses many proxies for accessibility:

1) The average accessibility index, that measures the average distance/time both in public a private transport to access from a given zone to all the others
2) The average weighted accessibility index, is the same that the previous one but weighted by the actual commuting flows between zones. This index accounts for the actual location of employment, and is plenty coherent with the bid rent theory, since it allows for account the residential/job location functional tensions
3) Finally, the author tests the distance to some “qualitative” hypothesis of centers such as the Plaça Catalunya, the Cinc d’Ors and the emblematic Plaça Francesc Macià.

His findings point an exiguous relationship between land value and accessibility indexes. On the contrary, the distance to “historic” landmark centralities are highly correlated to land prices. These results suggest that the trade-off mechanism behind transport costs between the site of residence and the site of employment exerts a weak influence
on residential prices. Conversely, social hierarchy imbedded in qualitative centralities is more correlated with real estate prices. In Roca (1986) the author studies price formation in the context of the 26 municipalities of the Metropolitan Corporation of Barcelona. Again, the results of this latter study, suggest that employment accessibility has a small role in explaining prices, that in some case become null or even erratic (i.e. prices increases as distance to employment centers increases). Such findings are compatible with the influence on prices that is exerted by other locative factors, such as the environment quality and the social segmentation produced by residential segregation that may imply both market premiums for prestigious locations and market discounts for stigmatized neighborhoods (Roca, 1988).

In the international scene, McMillen (1996) analyses the relationship between housing prices and the urban structure (proximity to CBD and subcenters) using historical data from Chicago finding an ambiguous relationship. Using a locally weighted regression the author estimates regression surfaces for the period 1936-1928 founding that the monocentric model correctly represents the spatial distribution of land values during the 19th century in such city. Nonetheless, in the post-war period, 1960-1990, especially after the suburbanization boom in the US the models suggest that Chicago is no longer a monocentric city, since subcenters, such as that located near the O’Hare Airport, are important employment clusters and as consequence also influence land prices. Such a finding was consistent with previously published evidence for the same city, as matter of fact McDonald and McMillen (1990) reported that, as early as 1928, that specific aeronautic subcenter was already significant in the distribution of land values. The same authors (McMillen & McDonald, 2002) also have studied whether land values were affected by the introduction of exclusive zoning for residential blocks in 1923 in Chicago. Their results indicate that exclusive zoning did not produce any modification of spatial trends in land values, except for that non-residential blocks where, departing up to zoning ordinance, were allowed to have also residence.

The problem with those studies is that most of them does not suffices in including all control variables that influence land values, such as amenities, externalities and social status, obscuring, in this way the role of centrality on real estate values. Aguirre and Marmolejo (2010) try to cope with these issues in their study for Metropolitan Barcelona, nevertheless they find multicollinearity problems since the distance to the nearest subcenter results pretty well correlated with the proximity to other subcenters and CBD. This study is different since it controls both the structural and location variables that may influence values, and non-distance to subcenter is used, but instead continuous indicators of centrality, as it is explained in the following section.

Case study, data and methodology

The study area is the Metropolitan Region of Barcelona comprised by 164 municipalities, nonetheless after discarding those municipalities without information or erratic outlier values only 137 municipalities containing 3,479 dwellings’ price offers are analyzed here.

Listing data was acquired from Habitaclia.com, as for November 2014, this website is one of the most important real estate listings specialized in sale and lease transactions in the residential market at Catalonia. The original data set retrieved comprises 35,116 selling cases for flats (multifamily dwellings), however, due the very recent obligation to obtain the energy certification in Spain (1st June 2013), a vast majority of cases did not report the energetic qualification claiming it was on pending of issue. Since this research was also concerned with the inclusion of such an environmental performance variable it was decide to exclude from the analysis all the cases where such information was not available.

Beyond asking prices, listing information includes variables regarding both the structural features of flats, and condominium services. Structural features include: area, number of rooms, number of baths, living room area, terrace/balcony area, story where the flat is placed, heat/air conditioning systems, information regarding renewal, energetic EPC class, penthouse position, number of levels (in the case of duplex/triplex dwellings), etc. Condominium services/features include: lift, swimming pool, private greenery, age, etc. Besides listing information, the following data and sources are used to control the effect of location on asking prices:

From the 2001 National Census, at census track level:
(Unfortunately, it was not possible to use 2011 data census since, due the crisis in Spain, such a census is based on a restricted sample survey, which is not reliable at census track level).

• Information regarding the social status of the neighborhood (e.g. education level, hierarchic position of resident employed population in the context or their productive organizations, percentage of residential buildings with doorkeeper service, etc.)
• Information regarding the accessibility level (e.g. declared time to get the workplace)
• Information regarding the environmental quality of the neighborhood (e.g. greenery perception)
Information regarding the available services (health, education, sociocultural, retail, office based services, etc.)

From the Cadastral database, at census track level:
- Built up surface, net built area, floor area allocated for different land uses

From the 2001 Metropolitan Mobility Survey, at transport zone level:
- Information regarding the chain-trip or chain of daily activities
- Information regarding the length and duration of travels

Using this latter information regarding the spatial-temporal behavior of people a synthetic indicator of centrality is built as follows:

1) Firstly, some intermediate-variables where computed. Some of such intermediate-variables are: time-density (Marmolejo & Cerda, 2012); diversity of activities; socioeconomic diversity of people that make activities in a given

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price (Euros)</td>
<td>3479</td>
<td>34.000</td>
<td>715.000</td>
<td>159.707</td>
<td>88.018</td>
</tr>
<tr>
<td>Unitary price (Euros/sqm)</td>
<td>3479</td>
<td>845</td>
<td>3,542</td>
<td>1,885</td>
<td>662</td>
</tr>
<tr>
<td>Area (sqm)</td>
<td>3479</td>
<td>25</td>
<td>234</td>
<td>84</td>
<td>28</td>
</tr>
<tr>
<td>Number of baths</td>
<td>3479</td>
<td>4</td>
<td>1,29</td>
<td>0,51</td>
<td></td>
</tr>
<tr>
<td>Number of rooms</td>
<td>3479</td>
<td>15</td>
<td>2,91</td>
<td>0,90</td>
<td></td>
</tr>
<tr>
<td>Baths/room</td>
<td>3459</td>
<td>2</td>
<td>0,48</td>
<td>0,23</td>
<td></td>
</tr>
<tr>
<td>Energy qualification (ordinal)</td>
<td>3479</td>
<td>7</td>
<td>5,29</td>
<td>1,25</td>
<td></td>
</tr>
<tr>
<td>Time-density</td>
<td>3479</td>
<td>-</td>
<td>13</td>
<td>2,14</td>
<td>1,63</td>
</tr>
<tr>
<td>Terrace/balcony area (sqm)</td>
<td>3479</td>
<td>-</td>
<td>256</td>
<td>9,73</td>
<td>14,53</td>
</tr>
<tr>
<td>Livingroom area (sqm)</td>
<td>3479</td>
<td>-</td>
<td>90</td>
<td>12,04</td>
<td>9,83</td>
</tr>
<tr>
<td>AC (dummy)</td>
<td>3479</td>
<td>-</td>
<td>1</td>
<td>0,29</td>
<td>0,46</td>
</tr>
<tr>
<td>Heating (dummy)</td>
<td>3479</td>
<td>-</td>
<td>1</td>
<td>0,42</td>
<td>0,49</td>
</tr>
<tr>
<td>Renewal (dummy)</td>
<td>3479</td>
<td>-</td>
<td>1</td>
<td>0,10</td>
<td>0,30</td>
</tr>
<tr>
<td>Penthouse (dummy)</td>
<td>3479</td>
<td>-</td>
<td>1</td>
<td>0,04</td>
<td>0,18</td>
</tr>
<tr>
<td>Duplex/triplex (dummy)</td>
<td>3479</td>
<td>-</td>
<td>1</td>
<td>0,06</td>
<td>0,23</td>
</tr>
</tbody>
</table>

| Structural features of Dwellings      |    |         |         |        |                |
| Structural features of Building       |    |         |         |        |                |
| Centrality                           |    |         |         |        |                |
| Built density (zone)                  | 3479 | 0,19    | 5,90    | 1,93   | 1,24           |
| Time-density                          | 3479 | 324     | 1,134,098 | 118,964 | 146,950       |
| Centrality index                      | 3479 | 2,52    | 20,41   | 11,29  | 2,29           |
| Diversity of land uses (zone)         | 3479 | 0,35    | 1,64    | 1,02   | 0,21           |
| Diversity of activities (zone)        | 3479 | 2,92    | 2,03    | 0,38   |                |
| Average time-to-work (zone)           | 3479 | 8,94    | 37,01   | 23,67  | 4,59           |
| Diversity of land uses at basament (zone) | 3479 | -     | 1,77    | 1,11   | 0,23           |

Notes:
Energy qualification 1=A=more efficient, 7=G=less efficient

Using this latter information regarding the spatial-temporal behavior of people a synthetic indicator of centrality is built as follows:

1) Firstly, some intermediate-variables where computed. Some of such intermediate-variables are: time-density (Marmolejo & Cerda, 2012); diversity of activities; socioeconomic diversity of people that make activities in a given
zone; distance travelled by people making activities in a given zone, etc. All the variables were computed for different
days in the week and 5 time-strips.

2) Secondly, intermediate-variables were encapsulated in a synthetic indicator of centrality using DP2 methodology

In Marmolejo & Cerda (2014) all the details concerning the construction of the synthetic indicator of centrality are
provided and theorized in the more general framework of time-geography, see in Figure 1 below the spatial distribution
of such a variable.

Since the geographical entities of data used are divergent: points for studied dwellings, census track for census and
cadastral data and transport zones for mobility information, it was used a geographic information system. Using a buffer
of 300 m (in addition, it was used a buffer of 600 m radius; nevertheless, the model built with such data was less able
to explain prices in comparison to that presented in this paper) radius around each flat and geospatial queries all the
information was transferred to the 35,116 selling cases contained in the original database. Table 1 summarizes basic
statistics for selected variables.

In Table 1 diversity indexes were computed using the entropy index proposed by Shannon:

\[ H_i = -1 \sum_{x} P_{x_i} \ln(P_{x_i}) \]  

In (1) \( H \) is the diversity index for an \( i \) zone, \( P \) is the probability to find the \( x \) category from \( n \) possible of a given
phenomenon in the \( i \) zone. (for example to find an office oriented premise among all the premises in zone \( i \)).

On the other hand, PCA factors refers to a principal component analysis carried out departing from the socioprofessional
classification of residents (managerial, professional, clerical, manufacturing, non-qualified, etc.). High values for PCA
factor for high professional profiles are associated to zones were live the working population occupying the most-
remunerated job positions. High values for PCA for low professional profile are associated with traditional zones of
residence of working-class population.

The final selection of the cases introduced in the model was the result of an elimination process as follows:

- Firstly, all those flats with unitary prices beyond +/-1 standard deviation from the average unitary price were
discarded.
- Next, a family of regression models was calculated, using the model with the best fit the Mahalanobis Distance
was computed. According to Marmolejo & González (2009) this procedure allows for the elimination of outliers
in the \( n \)-variables used in the regression analysis.
- Finally, using a sedimentation analysis it was detected the Mahalanobis Distance breaking point (i.e. the value
where the slope increases abruptly).
After applying the above stated process there were identified 3,479 valid cases.

The functional model used is log-lineal since it accomplishes with the basic statistics premises for OLS calibrating process: normality of residuals, homoscedasticity, and multicollinearity absence.

Results and discussion

The best of the models was able to explain 65.2% of the variance of asking prices of multifamily dwellings at the Metropolitan Region of Barcelona (see table 2).

Table 2. Model performance. Source: Own elaboration.

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>0.8081</td>
<td>0.653</td>
<td>0.652</td>
<td>0.28556</td>
</tr>
</tbody>
</table>

Table 3 reports the result of coefficients that have resulted significant at 95% of confidence; such coefficients are reported both in their original scale and in terms of z-values (Beta coefficients), this latter allows for comparison among independent variables. According to Beta coefficients the listing price is obviously influenced first by the floor area of the analyzed flats, nonetheless such a relation is not linear since the square of such area has a negative influence on prices. Such a finding suggests the existence of diminishing returns in the formation of prices, so the marginal addition of extra square meters produces a reduction in the willingness to pay in comparison to the previous incorporated area.

According to the beta coefficients, the next predicting variables are those related to the socioeconomic aspects of the neighborhood. In this dimension, the presence of resident population (PCA factor for high profile professions) employed in managerial and professional positions do have an important positive impact on real estate selling prices, as well as the presence of very well educated population in the surrounding of the dwelling (percentage of residents with university education). Also in this field it is worth to note the positive impact of the existence of buildings with high standing services such as doorkeeper. Whether those results evidence that such population has a higher purchase power or there is a market premium coming from the existence of a segregation process remains on the analysis.

Next, it appears an indicator of centrality (built density), which is associated to dense and central zones and proxies for the existence of local based services and land rent gradient explained both by accessibility and other amenities. It is worth to point that the synthetic centrality index, built on the spatial temporal behavior of population, also appears producing a complementary positive impact on residential asking prices. The inclusion of both of the continuous indicators of centrality in the model, having non-serious multicollinearity problems, means that they are measuring different aspects of centrality. Therefore, the built density (sq. meters of floor space/sq. meter of land) proxies for the historic fixed capital accumulated over the years; meanwhile the synthetic centrality index accounts for the actual use of the city by residents whose give contents and make differentiation between centers and non-centers.

Some structural variables associated to the quality of flats and buildings are in the next position with a positive impact such as the heat ventilation air conditioning, the interaction variable (Such interaction variable is constructed multiplying the level of the flat by a dummy variable, such a latter variable adopts the unitary value when the lift is present; and a negative unitary value otherwise) lift presence*story level where the flat is located, the existence of condominium services like swimming pool, the number of bathrooms. The energy qualification has a positive impact on prices, for each step that it increases, in the “ladder of environmental efficiency”, the listing prices increment 0.9%. That means that the difference on selling prices between the worst qualification (G) and the most efficient flats in terms of energy efficiency (A) is 5.11%, that implies that taking as a reference the mean value in the sample the difference is ceteris paribus 8.167 Euro.
Table 3. Model coefficients. Source: Own elaboration.

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
</tr>
<tr>
<td>14 (Constant)</td>
<td>10.297</td>
<td>0.055</td>
</tr>
<tr>
<td>Area (Sq. M.)</td>
<td>0.018</td>
<td>0.001</td>
</tr>
<tr>
<td>Residents with university degree (%)</td>
<td>0.005</td>
<td>0.001</td>
</tr>
<tr>
<td>AC</td>
<td>0.101</td>
<td>0.013</td>
</tr>
<tr>
<td>Area (Sq. M.)^2</td>
<td>-4.104E-05</td>
<td>0.000</td>
</tr>
<tr>
<td>Built density</td>
<td>0.038</td>
<td>0.006</td>
</tr>
<tr>
<td>Interaction lift * story level</td>
<td>0.013</td>
<td>0.002</td>
</tr>
<tr>
<td>Swimming pool</td>
<td>0.136</td>
<td>0.026</td>
</tr>
<tr>
<td>Number of baths</td>
<td>0.062</td>
<td>0.012</td>
</tr>
<tr>
<td>Heating</td>
<td>0.046</td>
<td>0.013</td>
</tr>
<tr>
<td>PCA factor for high professional profile</td>
<td>0.061</td>
<td>0.014</td>
</tr>
<tr>
<td>Residential buildings with doorkeeper (%)</td>
<td>0.005</td>
<td>0.001</td>
</tr>
<tr>
<td>Centrality synthetic index</td>
<td>0.010</td>
<td>0.003</td>
</tr>
<tr>
<td>Renewal</td>
<td>0.043</td>
<td>0.017</td>
</tr>
<tr>
<td>Energy qualification</td>
<td>0.009</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Figure 2 Analysis of residues. Source: Own elaboration.

Conclusions

Our findings suggest that, apart from structural variables concerning the quality and fittings of studied flats, centrality does exert an intermediate impact on real estate values. By using continuous centrality indexes such as that related to built density or the actual use of the city (synthetizing the spatial-temporal behavior of population) some of the habitual problems in multicentric hedonic price models have been solved with basic econometric approaches.

As it has been found previously in Barcelona (Roca, 1986, 1988, Roca et al., 2003, Marmolejo, 2008), most of the locative variables with the highest incidence on residential values are associate to the spatial segregation of population. So, locations characterized by very well educated population employed in managerial and professional positions do imply a significant and important positive impact on asking prices in our model. According to the b coefficient, which measures semi-elasticity in our log-lin model, by increasing 1% the neighbors with university studies, asking prices rise 0.5%, that
implies that the value gap between the zone with the lowest presence of high educated residents and the zone with the highest level of such population foster residential values 3.2 times (320%), by solely modifying this factor. Whether such finding is masking a higher acquisitive power and not or not only the existence of a market premium linked to the social perception (prestige/prejudice) remains in the impossibility to isolate the income effect since in Spain such data is not asked in the Census. It is necessary, thus, to review the enthroned importance of trade-off theory by relativizing it with other important factors such as socioresidential segregation. Especially in cities where real estate markets arbitrate both the formation of land values and land use allocation.

Acknowledgments

This paper has been produced in the framework of two research projects LugarES (MINECO CSO-2012-33441) and EnerVALOR (MINECO FEDER BIA2015-63606-R) leaded by the author.

References


