Journal of Housing Economics 29 (2015) 59–71

Contents lists available at ScienceDirect

Journal of Housing Economics

journal homepage: www.elsevier.com/locate/jhec

Use of a Gini index to examine housing price heterogeneity: A quantile approach ☆

Jaume García Villar [a], Josep Maria Raya [b,*]

[a] Department of Economics and Business, Universitat Pompeu Fabra, Spain

[b] Escola Superior de Ciències Socials i de la Empresa (Tecnocampus, Universitat Pompeu Fabra), C/Ernest Lluch 32, 08302 Mataro (Barcelona), Spain

ARTICLE INFO

Article history:
Received 1 June 2015
Accepted 8 June 2015
Available online 2 July 2015

JEL classification:
R21
C30

Keywords:
Quantile regressions
Housing prices
Hedonic models
Gini index
Heterogeneity in housing prices

ABSTRACT

This paper contributes to the existing literature that deals with the full distribution of house prices and its decomposition (primarily McMillen, 2008) by conducting a deeper analysis of housing price heterogeneity. Our approach differs from McMillen's insofar as our goal is to explain the variation in housing prices at a point in time rather than over a period of time. The basic statistic used to summarise house price distribution is the Gini index, which compares the actual distribution of the price per square metre (PPSM) with a uniform distribution. We decompose the Gini index into what can be explained for by the explanatory variables (which can also be easily decomposed into the contribution of each explanatory variable) and what remained unexplained. With a data set that includes appraisal values for 9297 dwellings in Barcelona in 1998–2001, the part explained by the standard OLS slopes (up to 60%) suggests a high degree of homogeneity in the linkage between PPSM and the explanatory variables. In any event, when heterogeneity is introduced using a quantile approach, that part of the Gini index explained for by the regressors falls. Finally, the variable that produces the most heterogeneity is area.

1. Introduction

Until recently, the housing economics literature has not dealt with the full distribution of house prices and specifically the decomposition of house prices. This paper contributes to the existing literature (primarily McMillen, 2008) by conducting a deeper analysis of heterogeneity of housing prices., McMillen, 2008) analysed changes in the full distribution of housing prices using data for sales of single-family dwellings in Chicago between 1995 and 2005. He found that the price of dwellings at higher percentiles rose faster than other housing and that the 2005 distribution implied a reduction in house price inequality. The main issue addressed by the paper is the relative contribution of the variation in quantile regression coefficients versus variation in the distribution of the variables explaining house price distribution. This paper complents McMillen (2008) 'is the ction section in order to emphasize the differences between the paper and the McMilnts McMillen (2008), who sets out to decompose changes in the distribution of house prices. Our approach differs from McMillen's insofar as our goal is to use Gini indexes to explain the variation in housing prices at a point in time rather than over a period of time. In this sense, unlike McMillen's, our paper is cross-sectional.

McMillen (2008), using a quantile approach, decomposes changes in the distribution of housing prices over time into items arising from: (1) changes in the distribution of the explanatory variables; (2) changes in the

☆ The authors would like to thank the Ministry of Technology and Science for the funding received through Project ECO2008-06395-C05-01.

* Corresponding author. Fax: +34 931696505.
E-mail address: josep.raya@upf.edu (J.M. Raya).

http://dx.doi.org/10.1016/j.jhe.2015.06.001
1051-1377/

distribution of estimated quantile coefficients of the hedonic price function. Quantile Regression (QR) is frequently used when estimation of the conditional mean cannot capture links between the dependent variable and the explanatory variables throughout the whole distribution of the former. The method is commonly used in other fields of economics, particularly in connection with inequality. QR has also recently been used in the literature on housing economics[1]. The approach used by McMillen is closely related to the literature on changes in earnings inequality. In particular, he follows Machado and Mata, 2005. Machado and Mata (2005), who propose a quantile regression-based decomposition based on the estimation of marginal wage distributions consistent with a conditional distribution estimated by quantile regression. Recently, Nicodemo and Raya (2012) present a relevant multi-city analysis.

In both health economics and labour economics, greater attention is paid to the decomposition issues addresses here. Wagstaff et al. (2002) demonstrate how the linear regression model can be used to decompose indices of inequality to identify the relative contributions made by the explanatory variables. This decomposition treats individuals' responses as homogeneous. Jones and Lopez (2002) demonstrate how this approach can be extended to allow individual heterogeneity through the use of quantile regression, producing an additional source of variation: the difference in coefficients across quantiles.

Our data span only four years (1998–2001) because our main aim is to provide a snapshot of the factors determining the housing price distribution during these years. The basic statistic used to summarise house price distribution is the Gini index (a universally used inequality index), which compares the actual distribution of the **price per square metre** (occasionally referred to here as **PPSM** for the sake of brevity) with a uniform distribution. We ranked housing prices from lowest to highest so we could compare the cumulative share with the 45-degree line distribution. We also studied inequality in the distribution of the PPSM variable and its decomposition into what can be explained for by the explanatory variables (which can also be easily decomposed into the contribution of each explanatory variable) and what remained unexplained. Our data set covered appraisal values for 9297 dwellings in Barcelona between 1998 and 2001. We used this to study changes in the Gini index over this short period of time and its decomposition into the portions due to the explained part and the unexplained part. Finally, we used a quantile model to capture information hidden in the unexplained part (unobserved heterogeneity) in order to estimate the full impact of the explanatory variables on the Gini index at any given point in the full price distribution.

The study is structured as follows: Section 2 demonstrates how the Gini index can be broken down into the contributions made by the various explanatory variables, whether these be individual, homogeneous or heterogeneous ones. We then present the empirical model by introducing the econometric methods that allow this decomposition. Section 3 presents the data and Section 4 the results. In the latter section, we also compare estimates based on OLS and on quantile regressions. Section 5 reports the Gini index findings and the impact of the explanatory variables on house prices using the previous estimates. Section 6 summarises the main conclusions of the study.

2. Methodology

The key methodology used in our study is the decomposition of a measure of inequality. We use the Gini index for price per square metre as a value for measuring housing inequality. The relative impacts of the explanatory variables on this measure are identified by means of a linear regression model.

The Gini index is an inequality measure that is usually associated with a descriptive approach to inequality measurement. The Gini index is named after the Italian statistician Corrado Gini, who invented the measure and published it in his 1912 paper *Variabilità e mutabilità.* It is closely linked to the representation of income inequality through the Lorenz curve[2]. In particular, it measures the ratio of the area between the Lorenz curve and the equidistribution line or 45-degree line (the extreme case where all incomes are equal) to the area of maximum concentration (the area between the Lorenz curve of an income distribution where all incomes are zero except for the last one and the equidistribution line).

This geometrical interpretation based on the Lorenz curve is, however, only one way to calculate the Gini index. Another approach, which proves particularly useful below, is to directly express the Gini index in terms of the covariance between the levels of a particular variables and its cumulative distribution (Lambert, 1993). In terms of our objective, we consider p_i to be the price per square metre of dwelling i and R_i the cumulative proportion of dwellings ordered according to p_i up to dwelling i (relative rank). Thus, the Gini index, G, for price per square metre is:

$$G = \left(\frac{2}{\bar{p}}\right)\mathrm{cov}(p_i, R_i) \tag{1}$$

where $\bar{p}$ is the average price per square metre and Cov is the covariance between the price per square metre p_i and the cumulative proportion of dwellings ordered according to p_i up to dwelling i.

Wagstaff et al. (2002) demonstrate how the linear regression model can be used to decompose indices of inequality so that one can identify the relative contributions made by the explanatory variables. Thus, following Wagstaff et al. (2002) if p_i is constructed from the following linear regression model (where ε_i are the residuals):

$$p_i = \beta_1 + \sum_{k=2}^{K}\beta_k x_{ki} + \varepsilon_i \tag{2}$$

[1] See McMillen and Thorsnes (2006), Coulson and McMillen (2007), Zietz et al. (2008) and Liao and Wang (2012).

[2] The Lorenz curve is a graphical representation of the cumulative distribution function of a probability distribution; it is a graph showing the proportion of the distribution taken up by the bottom y% of the values.

Substituting it with p_i, the Gini index of p may be written:

$$G = \sum_{k=2}^{K}\left(\beta_k \frac{\bar{x}_k}{\bar{p}}\right) C_k + \left(\frac{2}{\bar{p}}\right)\text{cov}(\varepsilon_i, R_i) \quad (3)$$

where the first term in brackets is the elasticity of p with regard to x_k evaluated in the sample mean and C_k is the concentration index[3] of x_k in p:

$$C_k = \frac{2}{\bar{x}_k}\text{cov}(x_{ik}, R_i) \quad (4)$$

Thus, following Eq. (3), the Gini index can be broken down into an "explained part" and an "unexplained part". The first term in (3) may be understood as the part explained by the regression. Furthermore, the explained part may easily be broken down into the contributions made by each explanatory variable.

The second term in Eq. (3) is the unexplained part, that is, the covariance of the term of the residuals (the portion of p_i which cannot be explained by x_k) in relation to the position of the individual in the distribution of the variable under consideration (R_i).

The unexplained part will be zero if the regression model for price per square metre is specified in such a way that there is no systematic variation in the unobserved heterogeneity of price per square metre in relation to the point occupied by the dwelling on the PPSM distribution. If this is the case, it makes no difference where we look on the price distribution – all explanatory variables weigh equally on the dwelling price. By contrast, the unobserved part of the decomposition of the Gini index is non-trivial if the unexplained part is not zero. In this case, there is unobserved heterogeneity and hence the relative impact of the explanatory variables on the Gini index depends on the price point on the full distribution.

Here, one should note that Jones and Lopez (2006) propose a method for dealing with this unobserved heterogeneity (the second term in Eq. (3)) while preserving useful information provided by the conventional regression model. Individual heterogeneity can generally be introduced into a regression model for price per square metre with heterogeneous parameters. Thus, following Jones and Lopez (2006), the regression model may be modified to yield the following result:

$$p_i = \beta_{i1} + \sum_{k=2}^{K}\beta_{ik}x_{ik} = X_i'\beta_i \quad (5)$$

where all of the parameters in the model are individual specific. It is worth noting that the intercept of this model β_{i1} includes both the individual systematic unobserved impact and purely non-systematic or random errors.

Substituting this last expression in (1), they obtain the following expression for the Gini index[4]:

$$G = \sum_{k=2}^{K}\beta_k^{MCO}\frac{\bar{x}_k}{\bar{p}}C_k + \left(\frac{2}{\bar{p}}\right)\sum_{k=2}^{K}\sum_{i}x_{ik}(\beta_{ik} - \beta_k^{MCO})(R_i - 1/2) + \left(\frac{2}{\bar{p}}\right)\text{cov}(\beta_{i1}, R_i) + \left(\frac{2}{\bar{p}}\right)\text{cov}(\varepsilon_i, R_i) \quad (6)$$

The first term of this equation is exactly the same as in (3), when the model (2) was estimated using OLS. The next two terms are the result of allowing heterogeneity. The second term in (3) is now divided into two parts, which become the following two terms in (6). The second term is the contribution to overall inequality of the covariance (weighted by the values of x_k) of the slope parameters with the relative rank. This part is the gain we obtain when we use a heterogeneous parameter model, since it is the part of the Gini index explained by heterogeneity. The third is simply the covariance of the intercept terms (individual systematic unobserved effect as well as random errors) with the relative rank, which is the only part that now remains unexplained.

In addition to the above, the decomposition of the Gini index contemplated in (3) allows us to identify the contribution made by each variable to the index. Obtaining this decomposition requires the linear regression model, which is used to characterise the distribution of the dependent variable according to its expectation value conditioned by the explanatory variables.

The regression model offers us estimates of the homogeneous parameters between individuals. Its application in the housing economics field is justified by hedonic price theory. According to this theory (Rosen, 1974), differentiated products are considered to possess various characteristics. The implicit marginal price of such a product may be ascertained through a hedonic price model that explains price it terms of those characteristics.

In the context of housing it is also easy to appreciate that the valuations individuals make of the physical characteristics of their dwelling or the area where they live differ depending on whether they have dwellings with a higher or lower price per square metre. It would therefore be interesting to know the behaviour of the explanatory variables over the PPSM distribution. For this, an estimator is required that allows heterogeneous responses. The one that fits the bill stems from the quantile regression (β_i)[5].

Summarising, the focus of the paper is to obtain the PPSM Gini inequality index for a sample of dwellings in Barcelona. Here, we wish to compare the full PPSM distribution with the equidistribution (where all the prices per square metre are equal). Furthermore, one can decompose this Gini index into various parts. The first part covers the explanatory variables (which can also be easily decomposed to yield the contribution made by each explanatory variable). The second part is the unexplained part, that is, the portion of p_i which cannot be explained by x_k. Using a quantile approach (heterogeneous parameter model) this part can be divided into two parts: (1) the contribution of the covariance between the coefficients of the explanatory variables in quantile regressions with the price per square

[3] This measure summarises inequality and is equal to the area between the Lorenz curve and the 45-degree line. The result is multiplied by 2 to ensure that the result is between −1 and 1.

[4] See Jones and Lopez (2002) for its complete derivation.

[5] The criterion of minimum absolute deviations tends to be used as described in Koenker and Bassett (1978).

metre (the second term in Eq. (6)); (2) the contribution of the covariance of the constants in the quantile regressions in relation to price per square metre (the third term in Expression (6)).

To obtain these results, we first estimated the hedonic price equation (2) by OLS and (5) by quantile regression. Second, we computed the Gini index with the OLS information (two parts) and with the quantile regression information (three parts). In both cases, we decomposed the information yielded into explanatory variables.

3. Data

The sample used in this study consisted of 9297 observations of Barcelona dwellings for the period 1998–2001. However, the sample did not constitute a data panel given that the dwellings in each annual sub-sample were different.

The data are from a representative annual sample collected by TINSA, an appraisal company. They contain a large variety of physical characteristics relating to housing, as well as information on the city area in which the dwelling is sited. The prices are therefore appraisal ones. These appraisals follow the criteria set out in Statutory Instrument ECO/805/2003 (market value) used for official and certain financial purposes (basically, mortgage valuation). The appraisal price is that at which a dwelling might be sold through a private contract struck between a buyer and seller on the appraisal date. One can therefore take appraisal prices as a surrogate for market prices – something that is supported by the fact that Spain's Ministry of Housing uses the former to calculate the housing price index.

Definitions of the variables employed in the study are as follows:

Price per square metre (PPSM): total value of the dwelling divided by the gross surface area (logarithmic scale).

Surface area: gross square metres including the corresponding proportion of communal areas.

Year: dummy variables defined according to the year each observation belongs to.

Age: seven dummy age variables in ascending order, for the following periods: new dwellings, from 1 to 5 years old, from 6 to 10, from 11 to 20, from 21 to 30, from 31 to 50, and finally over 50 years old.

Lift: dummy variable with value 1 or 0 according to whether the dwelling has a lift or not.

Floor: four categories were defined with their corresponding dummy variables according to whether the dwelling is on the ground floor or basement, first floor, second floor, or third floor and above.

Attic: dummy variable has a value of 1 if the dwelling is an attic dwelling and 0 if not.

Outward facing: dummy variable has a value of 1 if the property is outward facing and 0 if it is inward facing.

Heating: dummy variable has a value of 1 if the dwelling has heating and 0 if not.

Condition: what state the dwelling is in. Five dummy variables corresponding to each of the five appraisals of dwelling condition (very bad, bad, average, good and very good).

Renovation: time since last renovation. Four dummy variables corresponding to the four time periods defined for this variable: if the dwelling was renovated in the year the sample was taken or the five previous years, if it was renovated from 6 to 10 years previously, 11–20 or over 20.

Areas: statistical area (248) where the dwelling is located.

Table A.1 in the appendix shows the mean value of the characteristics (except for area) for each year of the sample. It will be seen that the sample consists mainly of used, recently renovated dwellings with a floor area of around 85 square metres.

Lastly, we should mention that both the data and the specification used in the estimates correspond to the model estimated in Garcia et al. (2006), which analysed the factors determining housing prices in the city of Barcelona. The main conclusions are that location accounts for over 50 per cent of the explained variation in prices and that education is the factor that best explains inter-area price variability. However, unlike in our paper, Garcia et al. (2006) do not address the full distribution and decomposition of housing prices.

In order to simplify the results, we substituted the dummy variables referring to location (areas) with the variable 'level of education'. This is justified by the fact that Garcia et al. (2006) demonstrate how variability related to the coefficients of the dummy variables referring to location (statistical area) is explained almost solely by the variable 'level of education' (an approximate correlation of 0.9). 'Level of education' is measured as the average number of years of study of the residents of an area. According to human capital theories and the empirical evidence available, this type of variable resembles the average income of residents of the area. As noted in Garcia et al. (2006), this strong correlation reveals a greater willingness to pay in locations where there is a higher level of education (or higher level of income) and could be explained by the fact that these variables (income or level of education) may correlate with other types of factors that individuals prize when choosing an area in which to live. To sum up, using 'level of education' rather than dummy variables for location means we can not only work with a much smaller number of variables but also capture the impact of location arising from the high correlation between the two variables. The phenomenon known as gentrification in the literature is based on a socio-economic classification of city areas (see Bridge, 1995). This approach is also supported by Gibbons and Machin (2003), who show that education is the key characteristic when evaluating the impact of location on housing prices".

4. Results

This section discusses the estimation of the 99 conditional quantile functions and the conditional mean function estimated by OLS. Table 1 shows the quantile coefficients of these variables for the quantiles 10, 25, 50, 75 and 90 (standard errors in brackets). Also shown are

Table 1
Quantile coefficients for percentiles 10, 25, 50, 75 and 90.

Percentile	10	25	50	75	90	OLS
Surface area	−5.571 (0.973)	−4.576 (0.678)	−5.126 (0.549)	−6.217 (0.618)	−5.947 (0.520)	−5.090 (0.398)
Surface area elasticity						
50 squared metre	−0.111	−0.092	−0.103	−0.124	−0.119	−0.102
100 squared metre	−0.056	−0.046	−0.051	−0.062	−0.060	−0.051
Heating	0.034 (0.006)	0.021 (0.006)	0.021 (0.004)	0.025 (0.004)	0.028 (0.006)	0.027 (0.004)
Outward facing	0.002[a] (0.007)	0.013 (0.004)	0.015 (0.002)	0.019 (0.005)	0.023 (0.006)	0.010 (0.004)
Lift	0.091 (0.006)	0.064 (0.005)	0.055 (0.003)	0.052 (0.007)	0.052 (0.008)	0.064 (0.004)
Year (Ref.: 1998)						
1999	0.256 (0.007)	0.249 (0.007)	0.239 (0.006)	0.238 (0.007)	0.242 (0.007)	0.251 (0.005)
2000	0.417 (0.009)	0.418 (0.006)	0.419 (0.004)	0.419 (0.003)	0.419 (0.006)	0.423 (0.004)
2001	0.576 (0.009)	0.581 (0.006)	0.579 (0.006)	0.575 (0.006)	0.582 (0.007)	0.585 (0.004)
Age (Ref.:>50)						
New	0.275 (0.024)	0.250 (0.012)	0.237 (0.008)	0.229 (0.014)	0.217 (0.017)	0.234 (0.012)
1 to 5 years old	0.141 (0.022)	0.128 (0.012)	0.133 (0.012)	0.130 (0.021)	0.092 (0.012)	0.117 (0.012)
6 to 10 years old	0.075 (0.011)	0.104 (0.017)	0.100 (0.017)	0.098 (0.014)	0.076 (0.013)	0.083 (0.012)
11 to 20 years old	0.094 (0.009)	0.089 (0.008)	0.080 (0.006)	0.069 (0.010)	0.051 (0.010)	0.078 (0.006)
21 to 30 years old	0.082 (0.009)	0.076 (0.007)	0.069 (0.006)	0.055 (0.007)	0.046 (0.007)	0.067 (0.005)
31 to 50 years old	0.038 (0.013)	0.035 (0.007)	0.027 (0.003)	0.020 (0.006)	0.008[a] (0.006)	0.026 (0.004)
Condition (Ref.: very bad)						
Bad	0.090 (0.021)	0.089 (0.010)	0.088 (0.007)	0.096 (0.013)	0.086 (0.012)	0.093 (0.008)
Average	0.179 (0.020)	0.166 (0.009)	0.159 (0.008)	0.161 (0.012)	0.153 (0.011)	0.172 (0.008)
Good	0.255 (0.023)	0.236 (0.011)	0.223 (0.010)	0.220 (0.013)	0.208 (0.013)	0.242 (0.009)
Very good	0.258 (0.031)	0.258 (0.014)	0.267 (0.008)	0.269 (0.015)	0.287 (0.021)	0.288 (0.013)
Floor (Ref.: ground floor)						
First	0.088 (0.024)	0.064 (0.016)	0.033 (0.010)	0.018 (0.009)	−0.003[a] (0.011)	0.037 (0.008)
Second	0.097 (0.021)	0.078 (0.016)	0.039 (0.012)	0.026[b] (0.008)	−0.004[a] (0.010)	0.044 (0.007)
Third	0.108 (0.020)	0.081 (0.013)	0.041 (0.011)	0.029 (0.008)	0.006[a] (0.010)	0.048 (0.007)
Attic	0.113 (0.025)	0.087 (0.015)	0.056 (0.016)	0.066 (0.014)	0.050 (0.020)	0.075 (0.010)
Lift*attic	0.015[a] (0.020)	0.030 (0.011)	0.026 (0.015)	0.005[a] (0.021)	−0.001[a] (0.026)	0.014[a] (0.011)
Renovation (Ref.:>20)						
0–5 years ago	0.072 (0.011)	0.058 (0.007)	0.069 (0.005)	0.068 (0.006)	0.072 (0.008)	0.073 (0.005)
6–10 years ago	0.069 (0.013)	0.045 (0.008)	0.054 (0.005)	0.045 (0.007)	0.057 (0.009)	0.060 (0.005)
11–20 years ago	0.050 (0.009)	0.032 (0.007)	0.035 (0.005)	0.024 (0.007)	0.029 (0.009)	0.038 (0.005)
Level of education	0.081 (0.002)	0.084 (0.002)	0.086 (0.001)	0.091 (0.002)	0.094 (0.002)	0.088 (0.001)
Constant	10.522 (0.032)	10.662 (0.030)	10.775 (0.021)	10.815 (0.023)	10.899 (0.028)	10.720 (0.017)

Note: standard errors in brackets.
[a] Not significant.
[b] Significant at 10%, the rest significant at 5% of the level of significance.

the coefficients of the explanatory variables of dwelling PPSM taken from the quantile regression and OLS estimates, together with the upper and lower ranges of the confidence interval at 95% of this latter estimate.

With regard to surface area, we adopt the specification $\ln\left(\frac{S}{1+S^{\theta}}\right)$, where θ is a parameter to consider and the elasticity (ε) of the price per square metre with respect to the surface area is $\varepsilon = \beta \cdot \frac{1+S^{\theta}\cdot(1-\theta)}{1+S^{\theta}}$, β being the coefficient of the variable $\ln\left(\frac{S}{1+S^{\theta}}\right)$. Note that the elasticity of the price per square metre (ε) is not constant, except in the case of $\theta = 0$, in which case the model presents a logarithmic specification for the area. In our case, for $\theta = 1$ it is possible to approximate the elasticity as $\varepsilon = \beta \cdot \frac{1}{1+S} \approx \frac{\beta}{S}$ and therefore β indicates the sign of the elasticity, whose magnitude depends on the value of the surface area, diminishing in absolute value as the area shrinks. One can also test for goodness of fit given that the proposed functional form includes the log functional form as a special case. The value of θ is estimated by maximum likelihood, we obtain a value close to 1, that is, we reject the model with constant elasticity ($\theta = 0$) and fail to reject the model with an elasticity where area $\theta = 1$. We therefore use a more flexible functional form than that usually used in the hedonic price literature (i.e. the log functional form, which implies constant price elasticity).

We observe how floor area has a more negative effect (coefficient) on price per square metre as we advance towards higher quantiles. In other words, in the OLS estimate (where the coefficient was constant and negative) the price sensitivity fell in absolute terms as floor area rose. That is, despite the fact that an increase in floor area reduces the price per square metre of the dwelling (note that this effect is less marked for larger dwellings).

In the case of quantile regression, we noted that β values did not remain constant and took values of between −4.58 and −6.22. However, with the exception of percentiles equal to or greater than approximately 0.75, these values are within the confidence intervals of the OLS estimate. The interpretation of this result is that at the higher percentiles, the effect of the same surface area increases (i.e. price per square metre falls more sharply).

An example will help to illustrate this result. The price elasticity of surface area of a dwelling measuring 50 square metres at percentile 25 is −0.0915, whilst the price elasticity per square metre for a dwelling with the same surface area at percentile 75 (where the minimum value of β is reached) is -0.124. However, the price elasticity per square metre for a dwelling measuring 100 m^2 at percentile 75 is -0.0621.

No clear pattern emerged regarding the impact that the time since the last renovation has on PPSM. Properties in the sample were generally renovated over twenty years ago. Only some of the values of the lower percentiles (10 and 25) lay outside the confidence interval of the OLS estimate.

The benchmark year used in the study was 1998. As far as time is concerned, the effect of each year on the dwelling's PPSM remains constant regardless of the percentile. In other words, after adjusting for quality, market price rises adds equal to the PPSM of all homes. In fact, the values of the quantile estimates remained within the OLS confidence interval but they almost always hovered near the lower limit.

With regard to the effect of the age of a dwelling on quantile regressions, it appears to have a diminishing quantitative impact as we advance towards higher percentiles. In most cases, estimates of the 10th and 90th percentiles do not appear within the confidence interval of the OLS estimate for the first and ninth deciles.

When compared to living on a lower floor, the price impact of living on higher floors or even in an attic flat also falls at higher percentiles. Quantile coefficients only appear within the OLS confidence interval in the central interval, which roughly runs from the first to the third quartiles.

The impact of having a lift remains constant from the 50th to the 90th percentiles, whereas it falls throughout the first 25 percentiles. Living in an attic flat increases the impact of a lift on the dwelling's PPSM. This effect does not show a clear pattern as we advance along the PPSM distribution, although the quantile estimates do remain within the confidence interval of the OLS estimates.

An outward-facing dwelling[6] pushes up the PPSM as one rises through the percentiles and is particularly marked at the top end of the range. The impact of heating on price follows a U-shape, falling from the 10th to the 50th percentile and then rising from the 50th to the 90th percentile. The coefficient estimates for the percentiles of dwellings with heating remain within the interval of the OLS estimate.

All other things being equal, a dwelling in better condition obviously commands a higher price per square metre – a finding that is constant across the percentiles. However, this variable has diminishing impact on price/m^2 for dwellings in average to good condition as one rises through the percentiles. By contrast, dwellings in very good condition rise in price/m^2 terms. For dwellings in bad condition, quantile estimates always appear within the OLS interval. In the case of dwellings in average, good and very good condition, these percentile estimates basically fall outside the interval of OLS estimates for the 90th percentile.

[6] Translator's note: Barcelona flats tend to be laid out around a central courtyard in the middle of city blocks. Outward facing flats give on to the street, while inward facing ones give on to the courtyard.

After obtaining the coefficient estimates for the variables of the five regressions for percentiles 10, 25, 50, 75 and 90, it is important to check, as Buchinsky (1998) stresses, whether they differ from one another statistically. The results of a quantile regression belong to what is known as a 'normal location model', providing the estimate of the constant is the only statistical difference between the percentiles. If this case, there is no heteroscedasticity in the model or put another way, the PPSM distribution is identical across all percentiles.

Table 2 presents statistical contrasts where the null hypothesis is that the slopes of the percentiles are equal for each of the variables. For example, we tested to see whether the impact of floor area on price per square metre is statistically equal (we cannot reject the null hypothesis that the coefficients of the percentiles are equal) in percentiles 10, 25, 50, 75 and 90. The table shows that the null hypothesis may only be rejected for location, lift, floor, heating, outward-facing dwellings, renovation 11–20 years earlier. Most of the dummy variables refer to the age of the property (except in the case of new dwellings and those from 6 to 10 years old).

To sum up, it would seem that the physical characteristics of dwellings display effects that are either statistically constant or that diminish at the higher end of the distribution (an exception here being the outward-facing variable). This decrease is offset by the intercept of and trend in the location variable. With regard to the intercept (which

Table 2
Slope equality test for quantile estimates for percentiles 10, 25, 50, 75 and 90.

	F (4 9270)	Prob > F
Surface area	2.26	0.06
Heating	2.53	0.04
Outward facing	2.60	0.03
Lift	8.19	0.00
Year (Ref.: 1998)		
1999	1.46	0.21
2000	0.02	1.00
2001	0.36	0.84
Age (Ref.:>50)		
New	1.35	0.25
1 to 6 years old	4.80	0.00
6 to 10 years old	1.55	0.19
11 to 20 years old	3.70	0.01
21 to 30 years old	5.61	0.00
31 to 50 years old	5.04	0.00
Condition (Ref.: very bad)		
Bad	0.32	0.86
Average	0.48	0.75
Good	1.11	0.35
Very good	0.62	0.65
Floor (Ref.: ground floor)		
First	3.29	0.01
Second	6.41	0.00
Third	6.63	0.00
Attic	2.00	0.09
Lift*attic	0.98	0.42
Renovation (Ref.:>20)		
0–5 years ago	1.06	0.37
6–10 years ago	2.60	0.37
11–20 years ago	3.39	0.01
Level of education	9.67	0.00

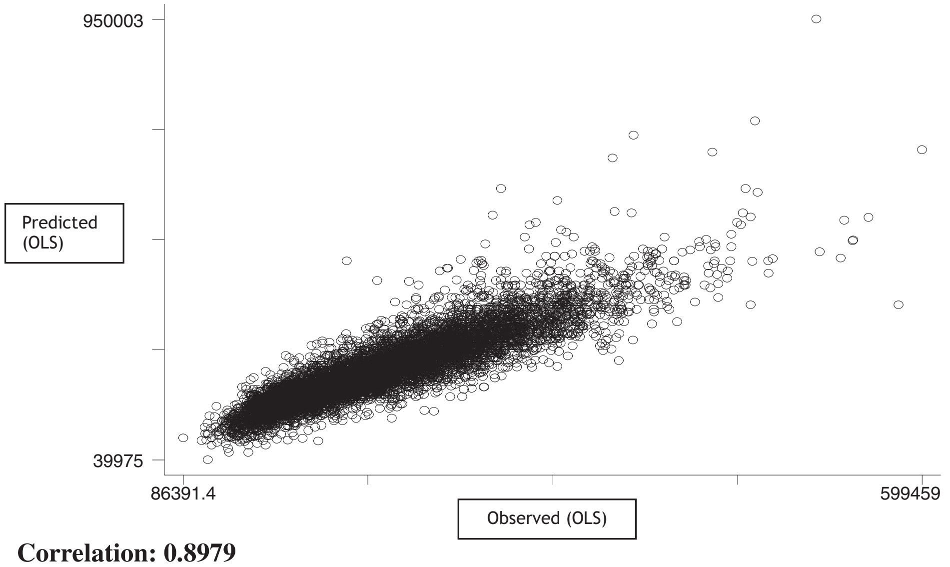


Fig. 1. Observed and predicted values for price per square metre in OLS estimate.

might be thought of as the price of a dwelling whose characteristics match the reference case), price/m^2 increases at higher percentiles. The most plausible explanation for the decreasing impact of physical characteristics in the upper part of the PPSM distribution is that at higher percentiles, physical characteristics count for far less than location. Thus, the effect of area[7] clearly increases as the percentile rises (see Table 1). The quantile coefficients only appear within the OLS confidence interval in the percentiles that are close to the mean. In Table 2, we see that there is rejection of the null hypothesis that the coefficients of the variable ‘area’ are equal across the percentiles.

The goodness of the procedure can be assessed by comparing the predictions made by quantile and OLS estimates. Figs. 1–6 display the relationship between the observed values and the predictions for the OLS and quantile estimates respectively, as well as reporting the value of the correlation coefficient between these variables. The quantile predictions were calculated by selecting the percentile prediction that yielded the lowest error between the observed and predicted values for each dwelling. The comparison clearly favours the heterogeneous parameter model. The correlation of the observed PPSM values versus those predicted by the model is 99% in the case of quantile regressions and 89% in the case of the OLS model. Table 3 shows that, on average, the quantile model errs by a tiny €0.14 when predicting a dwelling’s PPSM, whereas the comparable figure for the OLS model is €1.69. Nevertheless, the OLS model already includes a large part of the heterogeneity of the data, which is detected by higher R^2 around 80%. This can be explained by the data used (we have data for four years) and the goodness of the OLS procedure within the framework of hedonic pricing methodology. In order to clarify the extent to which time scale affects the data, Figs. 1–6 show the correlations of the predictions for the values observed for the OLS and quantile estimates from 1998 and 2001, respectively. The correlations of the observed values versus those predicted for the model range between 81% and 83% for the OLS model and 97% for the quantile regressions. Thus, although the R^2 of the OLS regressions drop to 66% and 69% for 1998 and 2001, they continue at very high levels, proving that the OLS model already captures much of the data heterogeneity, which in turn confirms the goodness of the hedonic methodology.

As we have seen, quantile regression helps us to capture all of the unobserved heterogeneity with higher R^2 (around 99%). Furthermore, it shows us the entire distribution of the coefficients, thereby adding to the comments made previously.

5. Analysis of the Gini index decomposition

Having got this far, we will now use the parameters estimated in the previous section to calculate and decompose the most commonly used inequality index in the literature: the Gini index[8]. The Gini index is calculated by comparing the distribution of the target variable (PPSM) with the equidistribution or 45-degree line (extreme case where all the prices per square metre are equal). The value of the Gini index will tend to zero as the corresponding distribution becomes more equal.

Tables 4a and 4b present the values of the Gini index for the years 1998 and 2001. The evidence points to differences that are of little significance for the two years studied. (the index falls slightly from 0.12374 to 0.10037). In any event, one can appreciate how the PPSM distribution becomes somewhat more equal (or, put another way, dwellings exhibiting lower PPSM in 1998 show greater price growth. This result is similar to that found in

[7] See Garcia et al. (2006), using the same data set, for more details regarding the result of distribution of areas in a city according to residents’ social and economic characteristics.

[8] Similar results are obtained when one conducts the exercise using the Theil index.

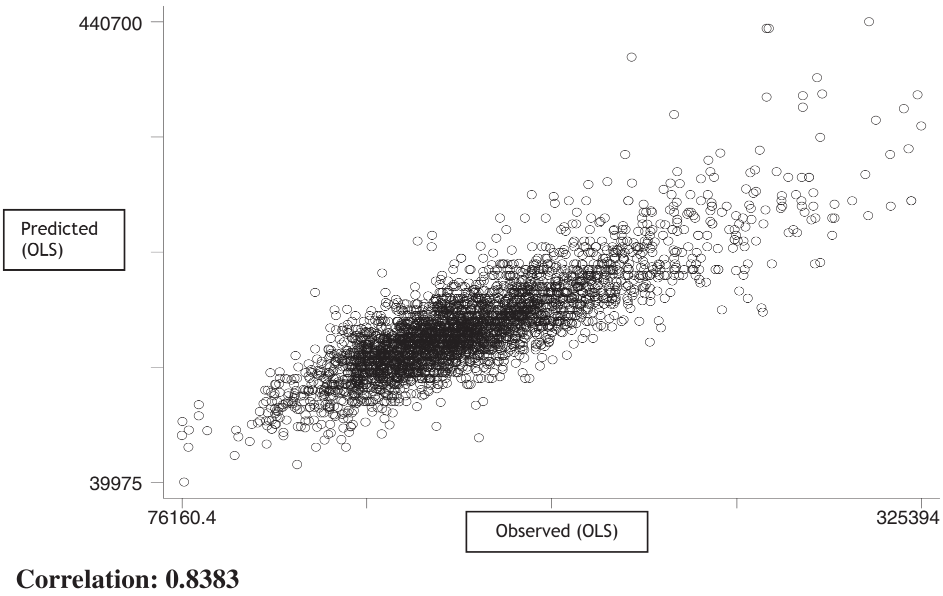


Fig. 2. Observed and predicted values for price per square metre in OLS estimate for 1998.

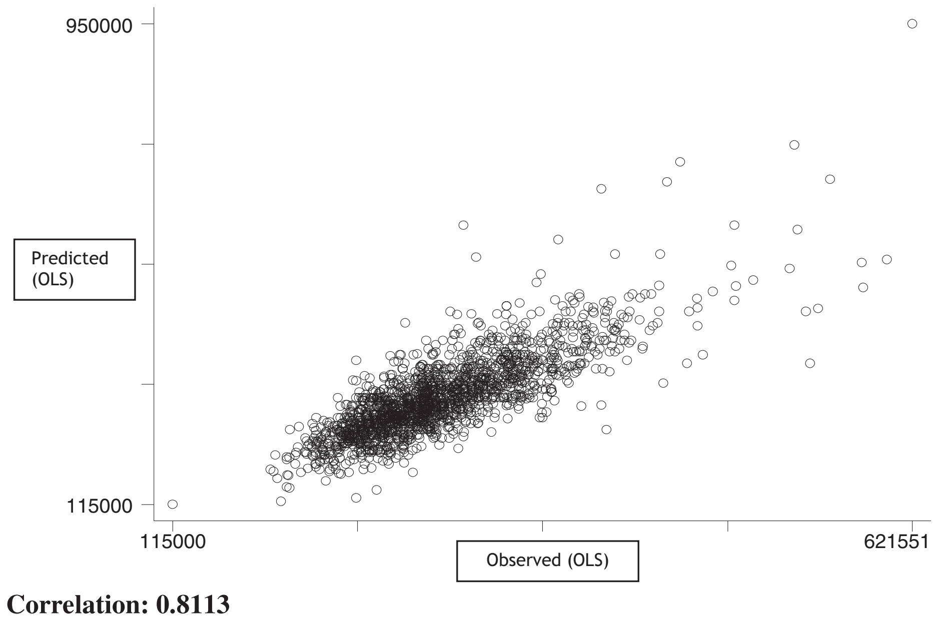
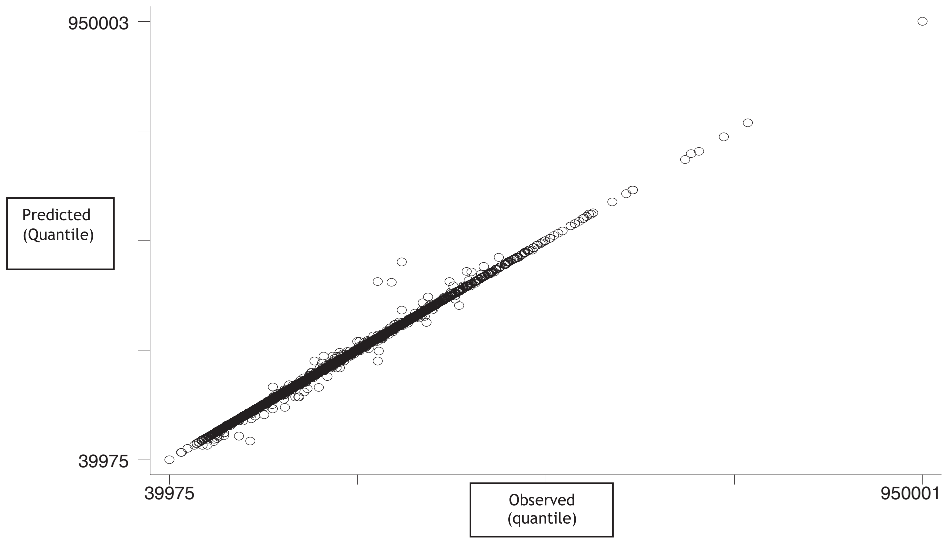


Fig. 3. Observed and predicted values for price per square metre in OLS estimate for 2001.

McMillen's paper (2008), in which the full distribution of the price in 2005 (versus the 1995 distribution) implies diminishing house price inequality. Both papers analyse data from periods of rapid appreciation. However, it should be pointed out that the values of the Gini index are low in both cases, indicating that the distribution of price/m^2 is very similar to the distribution where all dwellings have the same PPSM (an equidistribution).

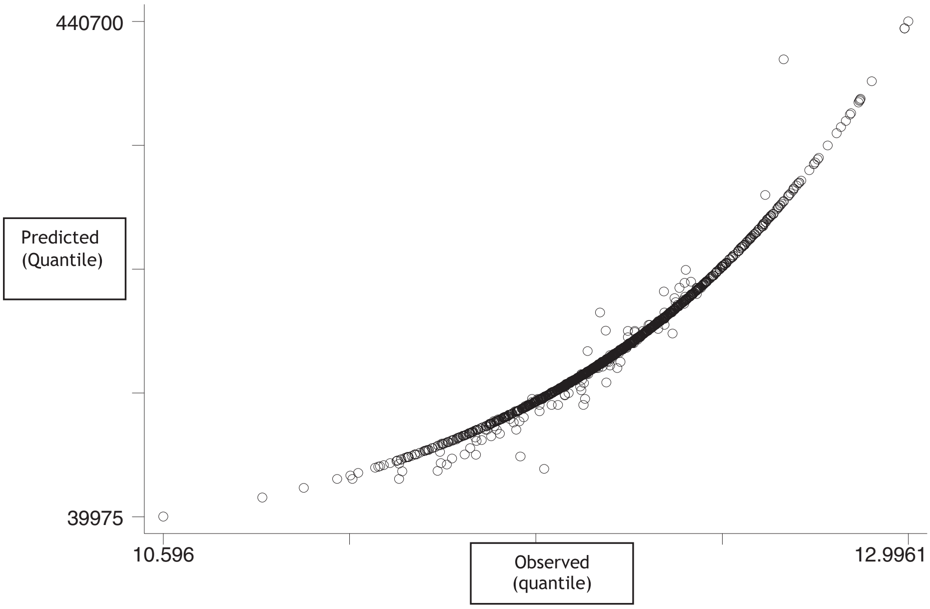
A decomposition of the Gini index for the first and final years of the sample (1998 and 2001, respectively) is given below with a view to analysing inequality and trends. Tables 4a and 4b show the results, which present this decomposition using both OLS estimates (first row) and quantile estimates (second row) by:

(a) The part explained by OLS. Contribution of the explanatory variables to the Gini index. Contribution of the product of OLS elasticities and the concentration indices of the explanatory variables in relation to price per square metre (first term



correlation= 0.9992

Fig. 4. Observed and predicted values for price per square metre in Quantile estimates.



correlation: 0.976

Fig. 5. Observed and predicted values for price per square metre in the Quantile estimates for 1998.

in expressions (3) and (6)). This part treats individuals as homogeneous since the effects of all the explanatory variables (β_k) are equal for all dwellings.

(b) The part unexplained by OLS and unobserved heterogeneity captured by the quantile coefficient estimates. Contribution of the quantile coefficient estimates to the Gini index. Contribution of the covariance between the coefficients of the explanatory variables in quantile regressions with the price per square metre (second term in Expression (6)).

(c) The part unexplained by OLS and unobserved heterogeneity of the quantile regression. Contribution of the quantile intercept to the Gini index. Contribution of the covariance of the constant in the quantile regressions in relation to price per square metre (third term in expression (6)). This is left unexplained by the OLS because we use the variation of the constant through the quantiles to furnish an explanation of this part of the Gini index.

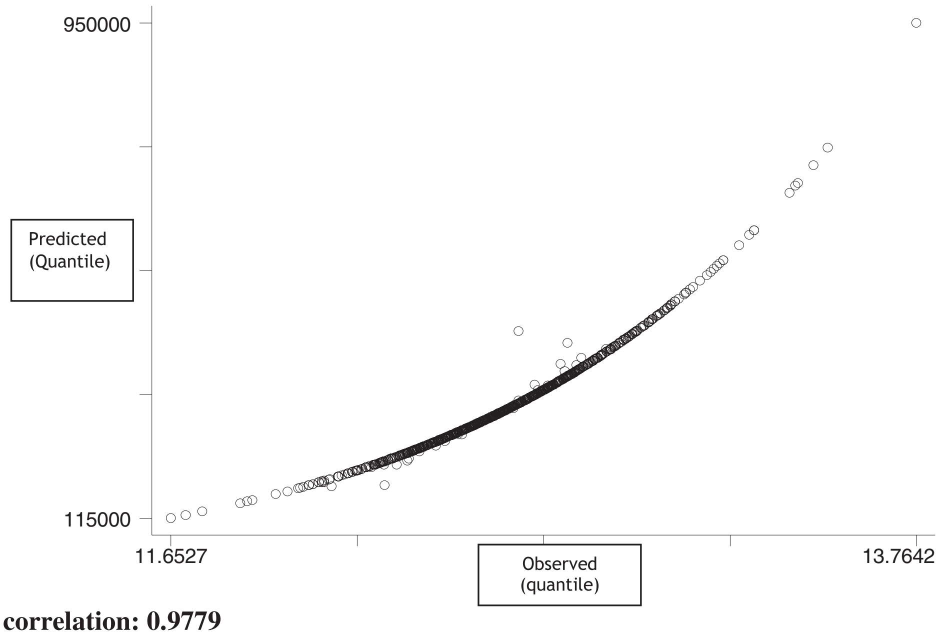


Fig. 6. Observed and predicted values for price per square metre in the Quantile estimates for 2001.

Table 3
Comparison of OLS and quantile predictive error.

Variable	Mean	Standard deviation
Price per square metre	1337.55	433.86
Absolute OLS error	11.61	190.99
Absolute quantile error	0.14	16.88

(d) The part unexplained by OLS. Contribution of the residual term to the Gini index. A residual that corresponds to the covariance of the prediction errors (of the OLS and the heterogeneous parameter model) in relation to the price per square metre (second term in expression (3) and last term in expression (6)).

Clearly, the value of the Gini index calculated by means of OLS decomposition Expression (3) has to be the same as that calculated by means of decomposing heterogeneous slopes Expression (6). The OLS decomposition treats the elasticities as homogeneous between individuals, just as they are treated in the estimate.

Firstly, the contribution of OLS slopes (the part explained by OLS and which treat individuals as homogeneous) are 60.35% and 64.86% for 1998 and 2001, respectively. These percentages are very high compared with the 22% obtained by Jones and Lopez (2006) in the context of health economics and are due to the goodness of fit shown by the OLS-estimated models used to explain the determinants of dwelling PPSM. In other words, over 60% of the inequality (Gini index) of the price/m^{2}can be explained by observed dwelling characteristics and under 40% remains unexplained.

These contributions demonstrate how variation in the explanatory variables influences the Gini index, maintaining parameters constant. Despite its explanatory power, the second term of the decomposition, which uses the heterogeneous parameter model (that is, the quantile estimates of the coefficients), reveals how the introduction of heterogeneity into the coefficients modifies the contribution of the explanatory variables. This part is the unobserved heterogeneity captured by the quantile coefficient estimates (e.g. 15.58% for 1998[9]). By taking the quantile coefficient estimates into account, we move towards a lower (i.e. more equal) Gini index. This shows how introducing heterogeneity into estimates of the model's slopes by means of quantile regression lowers the Gini index (in our case, by -9.25% and -7.30% for 1998 and 2001, respectively).

Tables 5a and 5b present the contributions of each of the explanatory variables to the Gini index for 1998 and 2001. The results group the variables into blocks. Thus, the total contribution of each of the variables to the Gini index can be described as follows. First, let us analyse the last column of Table 5a and Table 5b. A positive (negative) sign in the percentage of total contribution of one variable to the Gini index can be interpreted as an increase (decrease) in the Gini index, that is, a more unequal (equal) PPSM distribution. Despite this, the variable that contributes most to increasing the Gini index is area (level of education), which raises the index by 28.94% and 36.28% in 1998 and 2001, respectively. The variable 'lift'[10] increases the Gini index by 4.83% in 1998 and 8.24% in 2001, whilst heating does so by 3.45% and 4.07% in 1998 and 2001 respectively. Finally, non-consideration of the variable 'floor' would raise the index by 17.04% for

[9] This percentage comes from calculating the ratio 9.25/59.35, where 59.35 is the sum, in absolute values, of the percentage contribution of all three components of the unexplained part (9.25 + 49.50 + 0.60). This percentage is 14.68% in 2001.

[10] Translator's note: or "elevator" in American English.

Table 4a
Decomposition of the Gini index for 1998.

Index	OLS Contrib	Slopes %	Slopes Contrib	Quantile %	Constants Contrib	Quantile %	Residuals Contrib	%
Gini = 0.12374	0.07468	60.35					0.04906	39.65
Gini = 0.12374	0.07468	60.35	−0.01145	−9.25	0.06126	49.50	−0.00075	−0.60

Table 4b
Decomposition of the Gini index for 2001.

Index	OLS Contrib	Slopes %	Slopes Contrib	Quantile %	Constants Contrib	Quantile %	Residuals Contrib	%
Gini = 0.10037	0.06510	64.86					0.03527	35.14
Gini = 0.10037	0.06510	64.86	−0.00733	−7.30	0.04257	42.41	0.00002	0.02

Table 5a
Contribution of each explanatory variable in the decomposition of the Gini index for 1998.

Variable	OLS Contrib	%Contrib	Quant Contrib	%Contrib	Total Contrib	%Contrib
Surface area	−0.003107	−2.51%	0.009023	7.29%	0.005916	4.78%
Age	0.010922	8.84%	−0.001336	−1.09%	0.009585	7.74%
Lift	0.008155	6.59%	−0.002174	−1.76%	0.005981	4.83%
Floor	0.000429	0.35%	−0.021511	−17.38%	−0.021082	−17.04%
Lift*attic	0.000520	0.42%	0.000246	0.20%	0.000766	0.62%
Outward facing	0.000060	0.05%	0.002675	2.16%	0.002735	2.21%
Heating	0.004521	3.65%	−0.000249	−0.20%	0.004272	3.45%
Condition	0.010165	8.21%	−0.046844	−37.85%	−0.036679	−29.64%
Renovation	0.001239	1.01%	−0.002815	−2.27%	−0.001576	−1.28%
Area level of education	0.041783	33.75%	−0.005967	−4.86%	0.035811	28.94%
Total slopes	0.074688	60.35%	−0.011454	−9.26%	0.063234	51.10%
Constant			0.061264	49.50%	0.061264	49.50%
Total parameters					0.124498	100.60%
Residuals					−0.000749	−0.60%
Total					0.123749	100.00%

Table 5b
Contribution of each explanatory variable in the decomposition of the Gini index for 2001.

Variable	OLS Contrib	%Contrib	Quant Contrib	%Contrib	Total Contrib	%Contrib
Surface area	−0.003118	−3.11%	0.002679	2.67%	−0.000439	−0.44%
Age	0.004234	4.21%	−0.004116	−4.10%	0.000117	0.12%
Lift	0.009757	9.72%	−0.001487	−1.48%	0.008270	8.24%
Floor	0.000708	0.70%	−0.008462	−8.43%	−0.007754	−7.72%
Lift*attic	0.000568	0.57%	−0.000103	−0.10%	0.000465	0.46%
Outward facing	0.000009	0.01%	−0.000790	−0.79%	−0.000781	−0.78%
Heating	0.003669	3.66%	0.000413	0.41%	0.004082	4.07%
Condition	0.010639	10.60%	−0.009742	−9.71%	0.000897	0.89%
Renovation	0.000992	0.98%	−0.000210	−0.21%	0.000782	0.78%
Area level of education	0.037645	37.46%	−0.001249	−1.25%	0.036395	36.28%
Total slopes	0.065102	64.86%	−0.007328	−7.30%	0.057774	57.56%
Constant			0.042574	42.41%	0.042574	42.41%
Total parameters					0.100348	99.97%
Residuals					0.000020	0.03%
Total					0.100369	100.00%

1998 and approximately 7.72% for 2001, whilst non-consideration of the variable ‘condition’ would boost it by 29.64% in 1998 but only by 0.89% in 2001. Lastly, we should mention the increase in the Gini index in 1998 due to the variables ‘surface area’ (4.78%) and ‘age’ (7.74%).

To sum up, the variable that produces the most unequal distribution is “area” (level of education). The age of the property weighs less, though still heavily in 1998, as does having a lift (in 2001), producing a more unequal PPSM distribution. However, the floor (in both years) and the condition (in 1998) are the variables which help to produce a more equal PPSM distribution. Of all the variables, the one that makes the largest biggest contribution (in this case to inequality) is “area”. In other words, the PPSM dispersion increases when area is taken into consideration. The result arises from the high explanatory power of this variable and is not new in the literature. For instance Garcia et al. (2006) found that location explains 53.58% of total price variability. This result highlights a greater propensity to pay for properties located in better areas (i.e. ones where residents have higher levels of income, education or both). Moreover, higher education may be correlated with other characteristics that individuals value when taking housing decisions (parks, recreation areas, etc.).

Nevertheless, in the column showing the percentage of contribution of the quantile coefficients, most explanatory variables lead to a lowering of the Gini index (that is, making it more equal) when heterogeneity is introduced into the parameters. Thus, for 1998, with the exception of surface area (which leads to a 7.29% increase in the Gini index), the rest of the variables lead to a lowering of the Gini index if one considers that their effects vary throughout the PPSM distribution. The following cases stand out: condition (−37.85%), floor (−17.38%) and level of education (−4.86%). For 2001, with the notable exception of area, they all lower the Gini index: condition (−9.71%); floor (−8.43%); age (−4.10%).

The contribution of the covariance of the heterogeneous intercepts in relation to PPSM can be interpreted as the effect of unobserved heterogeneity, since the variation of the intercepts along the distribution of the price cannot be explained by any variable. If the explanatory variables were not related to the price per square metre in Tables 5a and 5b, the Gini index would increase through the contribution of the quantile constants: 49.50% and 42.41% of the current value for 1998 and 2001, respectively.

Lastly, the residual contributions to the Gini index for 1998 and 2001 when considering the heterogeneous parameters are practically zero. This contribution increases considerably in the OLS decomposition when the second and third terms in (6) are included.

6. Conclusions

This research studied the nature of the full distribution of house prices in the city of Barcelona using a sample of dwellings for the period 1998–2001. In particular, within the context of hedonic regression, we decompose the Gini index into the contribution of the explanatory variables to price per square metre (explained part) and the contribution of the residual (unexplained part). This unexplained part (unobserved heterogeneity) can be partially captured by the quantile coefficient estimates. The Gini index compares the actual price per square metre (PPSM) distribution with a uniform distribution. By ranking housing prices from lowest to highest, we can compare the cumulative share with the 45-degree line distribution.

In order to achieve this aim, we use OLS and a quantile approach to estimate a hedonic price model. In the quantile estimates we observe patterns of behaviour for some variables in relation to PPSM distribution in which the effect of the variables on higher quantiles generally decreases. Be this as it may, the effects between quantiles can be considered as statistically shifting for the following variables: floor, lift, number of years since last renovation, heating, age, outward facing and the variable that explains the effect of location (level of education). Here, valuation of the location (level of education) and the outward facing attribute of the property assumes greater importance at higher percentiles. It would therefore seem that these characteristics are taken into account more at the upper end of the distribution than at the lower one and hence can be considered luxury attributes. The opposite effect is observed, however, in characteristics such as age, floor and lift, which can be considered necessary attributes because they are taken into account more at the lower percentiles. Lastly, the effects of the dwelling having heating and the number of years since the last renovation display a non-monotonous pattern.

The Gini index fell slightly during the period 1998–2001, thereby making the PPSM distribution more uniform. It should be pointed out that the Gini index values are low in both years, indicating that the PPSM distribution is very similar to the one where all dwellings have the same price per square metre (an equidistribution). If we segregate by components, the explained part (contribution of the OLS slopes) increases from 60.35% to 64.86%, suggesting a high degree of homogeneity in the association of price per square metre with the explanatory variables. That is, considering the effect of each variable at the average point of distribution (as we do with the constant parameter model estimated through OLS) provides us with enough information to explain a large part of the Gini index. This is due to the fact that the hedonic price models estimated by means of OLS have a very high explanatory power (R^2 around 70%), which therefore explains a large part of the PPSM variability.

Heterogeneity can be introduced in order to estimate responses of the dependent variable to changes in the explanatory variables by means of the contributions of quantile regression slopes (the second component of the decomposition of the Gini index). Using quantile coefficient estimates explained some of the unobserved heterogeneity. For example, they revealed 15.58% of the part that was unexplained by OLS for the 1998 data. In any event, when heterogeneity was introduced using a quantile approach, the Gini index fell. This contribution to the lower value manifests itself significantly, in relation to its OLS contribution, in variables such as floor and condition.

In terms of the contribution of each explanatory variable, the one that produces the less uniform distribution is "area" (level of education). The effects of property age in 1998 and having a lift (in 2001) were less marked (though still important) in made for a less uniform PPSM distribution. However, "floor" (in both years) and "condition" (in 1998) are the variables that made for a more uniform PPSM distribution.

Knowing the full PPSM distribution would greatly improve housing price indexes to give a complete picture of the variation in appreciation rates (McMillen and Thorsnes, 2006). As far as heterogeneity is concerned, knowing the full PPSM distribution could enhance our knowledge of the Gini index and thus facilitate testing of heterogeneity trends when house prices are either rising or falling. Furthermore, decomposing the contribution of the explanatory variables would help scholars and governmental bodies discover which housing characteristics widen heterogeneity and which ones narrow it. An interesting avenue for future research would be to test whether the results presented in this paper can be extended to sales data (as opposed to appraisal data), periods of falling house prices (as opposed to rising house prices), and housing markets other than the Spanish one.

Appendix A.

Table A.1

Table A.1
Mean of the characteristics in the period 1998–2001.

	1998	1999	2000	2001
Surface area	86.02	83.97	84.42	85.86
Heating	0.326	0.353	0.351	0.625
Outward facing	0.829	0.855	0.846	0.827
Lift	0.642	0.605	0.590	0.590
Age				
New	0.032	0.030	0.043	0.034
1–6 years	0.026	0.016	0.020	0.026
6–10 years	0.020	0.016	0.023	0.013
11–20 years	0.171	0.126	0.087	0.073
21–30 years	0.327	0.316	0.292	0.260
31–50 years	0.208	0.246	0.279	0.294
Over 50 years	0.215	0.249	0.256	0.298
Condition				
Very bad	0.002	0.170	0.003	0.023
Bad	0.040	0.297	0.052	0.068
Average	0.613	0.367	0.639	0.614
Good	0.315	0.119	0.276	0.231
Very good	0.030	0.047	0.031	0.061
Renovation				
0–5 years	0.455	0.433	0.415	0.448
6–10 years	0.167	0.176	0.173	0.162
11–20 years	0.165	0.167	0.173	0.159
Over 20 years	0.180	0.194	0.196	0.188
Floor				
Ground	0.046	0.051	0.061	0.070
First	0.133	0.167	0.136	0.147
Second	0.157	0.174	0.177	0.180
Third or higher	0.607	0.497	0.508	0.545
Attic	0.055	0.109	0.115	0.055

References

Bridge, G., 1995. The space for class? On class analysis in the study of gentrification. Transac. Inst. Br. Geographr. New Series 20, 236–247.

Buchinsky, M., 1998. Recent advances in quantile regression models: a practical guideline for empirical research. J. Hum. Resour. 33, 88–126.

Couson, N., McMillen, D., 2007. The dynamics of intraurban quantile house price indexes. Urban Stud. 44 (8), 1517–1537.

Garcia, J., Matas, A., Raya, J.M., Raymond, J.L., 2006. Precios de la vivienda y localización: evidencia empírica para la ciudad de Barcelona. Papeles de Econo. Española 109, 125–139.

Gibbons, S., Machin, S., 2003. Valuing primary schools. J. Urban Econ. 53, 197–219.

Jones, A. and Lopez, A., 2002. The Importance of Individual Heterogeneity in the Decomposition of Measures of Socioeconomic Inequality in Health: An Approach Based on Quantile Regression, UPF Economics and Business Papers 626.

Jones, A., Lopez, A., 2006. Allowing for heterogeneity in the decomposition of measures of inequality in health. J. Econ. Inequal. 4 (3), 347–365.

Koenker, R., Bassett, G., 1978. Regression quantiles. Econometrica 46, 33–50.

Lambert, P., 1993. The distribution and redistribution of income: a mathematical analysis. Manchester University Press.

Liao, W., Wang, X., 2012. Hedonic house prices and spatial quantile regression. J. Hous. Econ. 21, 16–27.

Machado, J.A.F., Mata, J., 2005. Counterfactual decomposition of changes in wage distributions using quantile regression. J. Appl. Econometrics 20, 445–465.

McMillen, D., Thorsnes, P., 2006. Housing renovations and the quantile repeat sales-price index. Real State Econ. 34 (4), 567–584.

McMillen, D., 2008. Changes in the distribution of house prices over time: structural characteristics, neighborhood, or coefficients? J. Urban Econ. 64, 573–589.

Nicodemo, C., Raya, J., 2012. Change in the distribution of house prices across Spanish cities. Reg. Sci. Urban Econ. 42, 739–748.

Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition. J. Polit. Econ. 34–55.

Wagstaff, A., van Doorslaer, E., Watanabe, N., 2002. On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam. J. Econometrics 112, 207–223.

Zietz, J., Zietz, E., Sirmans, G., 2008. Determinants of house prices: a quantile regression approach. J. Real Estate Finance Econ, Springer 37 (4), 317–333.