Listing behaviour in the Italian real estate market

Rocco Curto, Elena Fregonara and Patrizia Semeraro
Department of Architecture and Design, Turin Polytechnic University, Turin, Italy

Received 29 January 2014
Revised 4 April 2014
Accepted 16 April 2014

Abstract

Purpose – The main purpose of this paper is to explore the listing behaviours of agents and sellers. In particular, the paper analyzes listing prices and the predicting power of the house features described in advertisements, to improve their use in real estate valuations. In Italy, selling prices are not public information and therefore listing prices play a key role for market analyses and are used by real estate companies and appraisers for estimating house values.

Design/methodology/approach – A traditional hedonic model was used to measure the overall contribution to listing price of the characteristics described in advertisements. The analysis was performed both on houses put on the market by agents and on houses put on the market by sellers. Listing price distributions and their deviation from normality were analyzed. Furthermore, a hedonic analysis was performed, which consisted of two steps. First, the coefficient of determination for any characteristic was computed. Second, the overall contribution to the listing price of the characteristics described in advertisements was measured.

Findings – The analysis shows the presence of factors which affect listing prices and which are not revealed to buyers in real estate advertisements. On the other hand, the presence of characteristics that do not affect the listing price but are described in advertisements was also found. Furthermore, agents and sellers showed different behaviours. While the marginal contributions of each characteristic estimated on a sample of houses put on the market by agents were significant, the analysis reveals that listing prices of houses put on the market by sellers are not explained by the house features.

Originality/value – To the best of the authors' knowledge, this is the first study to propose a hedonic approach to exploring the major determinants of listing prices of houses on sale on the Italian market. The listing behaviour of agents and sellers and the predicting power of the observable characteristics could address the use of listing prices in real estate valuations. At the same time, the potential presence of unobservable factors that affect the listing price could be a source of bias in estimating the value of houses.

Keywords Hedonic regression analysis, Listing prices

Paper type Research paper

1. Introduction

In Italy, selling prices of real estate are not public information and are difficult to observe. As a consequence, listing prices are used by institutional bodies to perform market analyses and by real estate companies and appraisers to estimate the value of houses. In this framework, listing prices represent an important signal of the house value. However, their importance is also recognized in the real estate literature; they are studied for their influence on the selling process. Several papers have empirically explored their impact on assets liquidity, measured by time on the market (Knight *et al.*, 1998) and price spreads, measured as the difference between the listing price and subsequent selling price (Song, 1995). Other papers, which investigated the influence of

International Journal of Housing Markets and Analysis
Vol. 8 No. 1, 2015
pp. 97-117

1753-8270
DOI 10.1108/IJHMA-01-2014-0003

listing prices on selling prices, found empirical evidence that selling prices are useful to improve selling price prediction (Knight, 2002; Horowitz, 1992).

As a result of the key role played by listing prices in Italy, we decided to explore listing prices by considering them as the main information available to estimate house value. Thus, we departed from the approaches described above and we built on the idea that listing prices are a function of house characteristics, combined with the seller's bargaining power, as empirically shown by Anglin *et al.* (2003). Therefore, we propose a hedonic approach (Rosen, 1974) to find the major determinants of listing prices and, consequently the house features to include in prediction models. Because of the hetereogeneity of assets, several explanatory variables could be included in the model. Thus, the choice of the characteristics becomes a key issue. Sirmans *et al.*'s (2005) review of the hedonic literature identifies the characteristics most frequently used to explain selling prices. Nevertheless, the contribution of house features to listing prices depends on agents' and sellers' knowledge of the contribution of characteristics to house value and on their listing strategy. In fact, the listing strategy consists in the choice of listing price (Beracha and Seiler, 2013), of the characteristics to include in the advertisement (Rodriguez and Siret, 2009), and of the language to be used to describe the property (Robertson and Doig, 2010), with the aim of selling the house as soon as possible and at the highest price (Yavas and Yang, 1995).

For our purposes, we decided to consider the house features most frequently described in advertisements, which are observable to consumers, buyers and appraisers. We built on the idea that the predicting power of observable house features could address the employment of listing prices in real estate valuations. We performed hedonic regression analyses to measure the contribution of individual observable characteristics to listing prices and a regression analysis to measure the overall contribution of the observable characteristics. The coefficients of determination of the models provide a measure of the price variation explained by the observable characteristics. Instead, the price variation which is not explained by the model exhibits the potential presence of unobservable determinants of listing prices.

Although real estate advertisements are usually put up by agents – under seller commitment – sellers could decide to put their house on the market by themselves. For this reason, we decided to perform two hedonic analyses:

(1) on a sample of houses put on the market by agents; and

(2) on a sample of houses put on the market by sellers.

The potential differences in agents' and sellers' use of house features to define listing prices may be a guide in the selection of the data to use in real estate valuations.

We analyzed a random sample size of 1,771 houses put up for sale in Turin – a city in the North of Italy – in the time period 2011-2012 on one of the most famous Italian websites for real estate sales.

The paper proceeds as follows: Section 2 introduces the methodology of analysis, Section 3 presents the empirical analysis and the statistical framework, Section 4 discusses the results and the conclusion follows.

2. The methodology

We used a traditional hedonic approach to measure the contribution to listing prices of the main features described in advertisements, which are observables to consumers,

appraisers and practitioners working in real estate companies. Unlike selling prices, listing prices are not the result of a negotiation process, but are defined by agents and sellers; thus, they reflect the heterogeneity of agents' and sellers' listing behaviours (Glower *et al.* (1996)). Indeed, we started our analysis by testing their distribution to find possible deviations from normality, a potential source of bias in estimates. As a consequence, when necessary, we adopted a parametrical transformation of listing prices to reduce the deviation from normality and improve the fit of the regression model. In particular, we considered the Box-Cox transformation, which is recalled in the Appendix, and the traditional logarithmic transformation, which is a particular Box-Cox transformation. The explanatory variables of the model were defined using the observable characteristics described in real estate advertisements. In particular, we identified the main features (including location) described in the real estate advertisements of one of the most famous real estate websites in Italy, from which the data were selected.

The explanatory variables included in hedonic models are traditionally grouped into house features and location amenities. Because we analyzed an Italian city, where most of the houses are, we grouped house features into building characteristics and apartment characteristics. The building and apartment characteristics described in advertisements may be modelled by introducing dummy variables. Instead, location is revealed to buyers by means of the house address, and sometimes, with a brief description of the amenities in the area. At present, many papers deal with the introduction of spatial statistics to model the spatial effect on the house value (Pace *et al.*, 1998) for an overview on spatial statistics in the real estate literature). We depart from this approach, as we analyze the behaviour of sellers and agents that are usually not familiar with advanced statistics. Some recent papers (Goodman and Tibodeau, 1998 and Bourassa *et al.*, 2007) conclude that geographical housing submarkets, such as the ones defined by agents, are more important in predicting house prices than the spatial statistics approach (Bourassa *et al.*, 2003 and Bourassa *et al.*, 2008). The determination of housing submarkets requires an analysis of the spatial market structure (Kauko, 2006), which is outside the aim of the present paper. Indeed, we decided to include location in the model by using a geographical segmentation which is typical of the Italian market and seems to us a reasonable benchmark for both agents and sellers.

Once we had defined the observable explanatory variables, we performed two steps of work. The first step aimed at isolating the contribution of each individual characteristic to listing price. We empirically computed the coefficient of determination corresponding to each explanatory variable in the following regression model:

$$Y_{LP} = \alpha_0 + \sum_{j=1}^{n} \alpha_j X_j + \varepsilon \tag{2.1}$$

where Y_{LP} is a suitable transformation of the listing price, X_j the dummy variables defined by the characteristic $n + 1$ levels (X_0 is the omitted level), the hedonic weight α_i, i = 1 […], n assigned to each variable is equivalent to the corresponding dummy's overall contribution to the value of price (Rosen, 1974), α_0 is the model intercept and ε the error term. By so doing, we identified the house features which had a negligible contribution to price and that were only used to promote the house and to attract potential buyers. The second step of analysis introduced a hedonic model to measure the

overall contribution of the main observable features on the listing price. The hedonic model takes the following form:

$$Y_{LP} = \alpha_0 + \sum_{j=1}^{n_1} \alpha_{1j} X_{1j} + \ldots + \sum_{j=1}^{n_N} \alpha_{Nj} X_{Nj} + \varepsilon \quad (2.2)$$

where, Y_{LP} is a suitable transformation of listing price, the variables X_{ij}, $j = 1 [\ldots] n_N$, $i = 1 [\ldots]$, N are the dummy variables introduced for any of the N observable characteristic, the hedonic weight α_{ij}, $j = 1 [\ldots] n_i$, $i = 1 [\ldots]$, n assigned to each dummy variable is equivalent to this characteristic's level contribution to the value of price (Rosen, 1974), α_0 is the model intercept and ε the error term. The hedonic weights in equations (2.1) and (2.2) were estimated using traditional least squares estimates. The coefficient of determination of the regression model (R^2) measures the proportion of variation of the dependent variable (transformed listing price) explained by the model. Each step was performed on a sample of houses put on the market by agents and on a sample houses put on the market by sellers. The comparison of the results highlights the differences between the behaviour of agents and sellers in regard to the contribution of house features to listing price. The following section presents the case study; it introduces the main characteristics included in Internet advertisements, i.e. the explanatory variables of the regression model, and presents the sample statistics.

3. Empirical analysis

3.1 Data

We examined a random sample of 1,177 Internet real estate advertisements of houses on sale in the city of Turin in the time period of 2011-2012. The data were sampled from one of the main Italian real estate websites.

The sample belongs to the databases of the Turin Real Estate Market Observatory (TREMO). TREMO was founded in 2000 through an agreement between the *Politecnico di Torino*, the Municipality and the Chamber of Commerce.

Before introducing the sample statistics, we list the set of the explanatory statistical variables included in the hedonic model in Table I, which includes their description and specifies the types of variables. The characteristics are grouped in apartment characteristics, building characteristics and location. As outlined above, location is specified by dummy variables representing geographical submarkets provided by the Italian law[1]. In Italy, every city has to be divided into homogeneous cadastral zones, named Microzones. In this respect, Italian real estate observatories publicize house price (usually listing prices) statistics for each Microzone. As a consequence, descriptive statistics of listing prices for each Microzone are available to sellers, buyers and appraisers.

The city of Turin was divided into 40 Microzones by the Politecnico di Torino, according to the law in 1999. Microzones are numbered from 1 to 40, fanning out from the centre of the city to the suburbs. The most dynamic are the semi-central submarkets – Microzones 9, 11, 15, 17, 29, 31 and 32 – built after 1960. Central Microzones are characterized by the presence of historical buildings and desirable properties that are often not listed in real estate advertisements. In particular, houses on sale in Microzone 16, which is a central pedestrian zone with attractive properties, are not promoted in

public advertisements. The hill zone, Microzone 24, is characterized by the presence of detached houses as well as apartment blocks. Note that the attractiveness of each Microzone depends on the subjective taste of sellers and buyers, which may prefer the city centre, the hill zone or the suburbs, even though they belong to the same socio-demographic group and have similar information about the amenities of the area (Kauko, 2006).

3.1.1 Remark 3.1 Before going into the empirical investigation, we note that the building condition is not included in the list of Table I. In fact, the building condition is usually not described in advertisements. Obviously, some pictures of the front of the building are included, but they do not allow buyers to infer the condition of the whole building. Nevertheless, several empirical results support the importance of the house maintenance level to explain selling price, as house repair costs are included in the house selling price (Knight and Sirmans, 1996; Knight *et al.*, 2000).

Obviously, the advertisements also include a reference to the real estate agency or to the seller who put the house on sale.

Type of characteristic	Characteristic	Type of variable	Description
	Microzone	Dummy	Geographical segmentation provided by Italian law
Apartment characteristics	Number of rooms	Numerical	
	Size	Continuous	Measured in *commercial* m^2, defined by the Italian law. The commercial size considers habitable areas, uninhabitable areas, size of balconies and size of terraces (large balconies)
	Floor	Dummy	
	Apartment condition	Dummy	The apartment condition is not described by agents, but it can be inferred from apartment pictures. Apartment condition information is expressed in three levels and is based on the apartment photos in the photo gallery attached to advertisements
	Number of bathrooms	Numerical	
Building characteristics	Building number of floors	Numerical	
	Building quality	Dummy	There are five building quality levels and are based on a cadastral definition and considers several building characteristics, such as the building materials and age of the building. The highest level corresponds to attractive properties while the lowest corresponds to council houses
	Presence of caretaker	Binary	
	Presence of a garage	Binary	

Table I.
House characteristics

3.2 The sample: descriptive statistics

This section analyses the whole sample and two subsamples that we named *Agents* and *Sellers*, used to perform the hedonic analysis. The sample *Agents* includes the houses put on sale by real estate agents, the sample *Sellers* the houses put on sale by sellers. The sample sizes are respectively 1,118 and 59. The proportion of observations in the sample *Sellers* (0.05) highlights that most of houses are put up for sale by agents. Descriptive statistics of the sample *Agents* and of the sample *Sellers* are provided in Table II. Descriptive statistics of the variable Microzone are provided in Tables III and IV. For completeness, we also provide descriptive statistics of the variable listing price (*LP*) on each Microzone in Tables II and III.

Notice that, due to the sample size of the sample *Sellers*, not all the Microzones are represented in the sample *Sellers*. As a consequence, we cannot estimate the marginal contribution of the missing Microzones to price.

Before performing the hedonic regression analysis, we computed the normalized concentration Gini index (*G*) for each characteristic to compare the distributions of the individual characteristics in the two samples (Table V). The definition of the normalized concentration Gini index is included in the Appendix. The coefficient is computed to underline that, although the sample *Sellers* is small with respect to the sample *Agents*, the characteristics have similar distributions – measured by their hetereogeneity – according to the fact that the samples are randomly selected. In fact, we notice that – except for the variable Agency – the two samples have similar measures of heterogeneity for each characteristic. Notice that the variable Agency in the sample *Agents* has a high Gini index ($G = 0.99$), showing that many different agencies use this selected website to put houses on sale. Thus, the present study does not represent the behaviour of a sole real estate agency.

4. Results: hedonic analysis

We started the empirical analysis by testing the distributions of listing prices in the three samples: whole sample, *Agents* subsample and *Sellers* subsample. We tested whether the three samples come from normal distributions, by performing the Shapiro-Wilk test. We analysed both listing prices and their logarithms, used in the hedonic regression analysis below. The Shapiro-Wilk test rejected normality of distributions of prices and log prices (p-values < 0.01) of the three samples. The following Table VI shows the empirical skewness s and kurtosis k computed from the three samples for listing prices *LP*, and their logarithms. As expected, listing price distributions exhibit fat tails and asymmetry, which confirm their deviation to normality. Notice that kurtosis is higher in the whole sample and in the sample *Agents*, which shows very fat tails. Obviously deviation from normality is smaller if we consider log-prices. Nevertheless, we also performed the regression analyses by applying the Box-Cox transformation to listing prices. The Box-Cox transformation parameter was estimated using maximum likelihood estimation. The estimated parameter was close to $\gamma = 0.05$. By applying the Box-Cox transformation to listing prices, i.e. $Y_{LP} = LP^{0.05}$, the coefficients of determination and the significance levels of the coefficients estimated are analogous to the results obtained using the logarithmic transformation. For this reason, we present the result obtained by estimating the models in equations (2.1)

Table II.
Samples statistics

Variable	DB agencies Mean	SD	Levels	Frequency	DB sellers Mean	SD	Frequency
Apartment							
LP	3,32,120.93	3,08,707.93			2,78,881.36	2,10,250.92	
Size	107.64	61.54			96.08	46.45	
Number of rooms	3.59	1.47			3.46	1.29	
Number of bathrooms	1.45	0.65			1.36	0.64	
Number of terraces	0.17	0.44			0.15	0.52	
Apartment condition (1 = worst, 3 = best)			1	0.06			0.00
			2	0.25			0.12
			3	0.47			0.54
			4	0.22			0.34
Floor			0	0.10			0.02
			1	0.24			0.20
			2	0.19			0.15
			3	0.15			0.20
			4	0.13			0.19
			5	0.08			0.14
			6	0.05			0.03
			7	0.03			0.02
			8	0.02			0.03
			9	0.01			0.02
Building							
Building number of floors	4.97	2.07			4.93	2.27	
Elevator (1 = presence 0 = absence)			0	0.33			0.31
			1	0.67			0.69
Presence of caretaker (1 = presence 0 = absence)			0	0.91			0.88
			1	0.09			0.12
Building quality (1 = worst, 5 = best)			1	0.03			0.07
			2	0.20			0.22
			3	0.51			0.44
			4	0.22			0.24
			5	0.04			0.03
Garage (1 = presence 0 = absence)			0	0.86			0.86
			1	0.14			0.14

and (2.2) using the logarithms of listing prices as dependent variables, i.e. $Y_{LP} = \log(LP)$, which is the transformation of prices most frequently used in the real estate literature.

Because of the samples heterogeneity and to the differences in *Agents* and *Sellers* sample sizes, we performed the analysis of regression residuals, to enforce our results and to highlight potential patterns in data. Residuals distributions are discussed at the end of the section.

Microzone	Frequency	Mean	SD
1	0.01	6,124.65	1,075.12
2	0.04	3,904.02	1,000.92
3	0.02	4,275.2	591.51
4	0.02	3,895.5	549.34
5	0.03	3,847.76	735.83
6	0.01	4,731.58	442.04
7	0.03	3,055.27	787.17
8	0.02	4,101.29	1,802.78
9	0.02	3,465.26	653.1
10	0.02	2,693.54	573.26
11	0.03	2,745.79	676.83
12	0.02	2,875.39	501.65
13	0	4,545.92	1,364.76
14	0.02	3,966.03	696.24
15	0.02	3,359.28	470.33
16	0.01	6,213.49	1,169.06
17	0.01	3,292.25	600.81
18	0.02	3,151.07	927.6
19	0.04	2,137.51	565.01
20	0.02	2,211.73	726.46
21	0.06	1,576.22	358.28
22	0.01	3,148.2	627.25
23	0.02	4,603.94	1,160.93
24	0.05	3,587.66	1,103.02
25	0.01	3,032.25	1,139.21
26	0.03	2,265.85	595.57
27	0.02	2,274.42	534.63
28	0.02	2,066.71	444.97
29	0.08	2,423.74	519.08
30	0.02	1,908.32	481.94
31	0.03	2,255.41	432.38
32	0.03	2,753.96	700.39
33	0.05	2,296.73	482.74
34	0.02	2,391.03	647.55
35	0.05	1,858.11	471.54
36	0.02	1,647.98	468.43
37	0.03	1,826.92	419.21
38	0.03	1,581.45	309.69
39	0.01	3,816.5	522.16
40	0.02	2,032.69	499.44

Table III.
Microzones sample statistics-sample agents

As a first step of the hedonic analysis, we performed a simple regression model to explain log-listing prices for each observable characteristic. For each regression, the adjusted coefficient of determination R^2_{adj} measured the listing price variation explained by the characteristic used in the model. The empirical coefficient of determination estimated on the three samples for each characteristic are in Table VII.

Let us consider the whole sample first. The empirical coefficients of determination R^2_{adj} (all significant except the coefficient of floor) indicate the presence of observable

Microzone	Frequency	Mean	SD
1	0.02	3,829.79	NA
2	0.00	NA	NA
3	0.03	5,944.12	503.29
4	0.00	NA	NA
5	0.02	4,285.71	NA
6	0.00	NA	NA
7	0.07	2,720.72	864.29
8	0.00	NA	NA
9	0.02	2,700.73	NA
10	0.02	2,680.00	NA
11	0.03	3,428.57	808.12
12	0.03	3,840.64	413.93
13	0.00	NA	NA
14	0.00	NA	NA
15	0.05	3,570.04	349.95
16	0.00	NA	NA
17	0.00	NA	NA
18	0.02	3,071.43	NA
19	0.03	1,958.79	159.29
20	0.03	1,651.05	63.07
21	0.02	2,222.22	NA
22	0.05	2,361.22	131.90
23	0.00	NA	NA
24	0.08	4,227.21	694.73
25	0.02	3,225.00	NA
26	0.00	NA	NA
27	0.02	3,250.00	NA
28	0.05	2,464.37	776.28
29	0.08	2,507.41	528.81
30	0.02	2,400.00	NA
31	0.08	2,215.28	320.03
32	0.03	2,545.49	133.66
33	0.02	1,904.00	NA
34	0.00	NA	NA
35	0.07	1,937.12	563.47
36	0.00	NA	NA
37	0.00	NA	NA
38	0.08	1,696.70	332.83
39	0.00	NA	NA
40	0.00	NA	NA

Table IV.
Microzones sample statistics-sample sellers

Table V.
Gini coefficient (G)

Characteristic	*G* (DB)	*G* (*Agents*)	*G* (*Sellers*)
Agency	0.99	0.99	0.00
Unit condition	0.89	0.89	0.77
Floor	0.93	0.93	0.92
Elevator	0.88	0.88	0.85
Presence of caretaker	0.33	0.32	0.42
Building quality	0.81	0.81	0.87
Box	0.48	0.49	0.47
Microzone	0.99	0.99	0.97
Year	1.00	1.00	0.61

Table VI.
Empirical skewness and kurstosis

Variable	*k*	*S*	*k_Agents*	*s_Agents*	*k_Sellers*	*s_Sellers*
Log(P)	2.67	0.32	2.64	0.31	3.16	0.62
p	15.74	2.83	15.57	2.82	6.74	2.05

Table VII.
Empirical coefficient of determination for each characteristic

Characteristic	Whole sample R^2 Adjusted significance	Sample agents R^2 Adjusted significance	Sample sellers R^2_{Adj} significance
Size	0.692***	0.693***	0.666***
Number of rooms	0.563***	0.566***	0.483***
Floor	0.003	0.003	−0.112
Unit condition	0.019***	0.046***	0.026
Number of bathroom	0.501***	0.504***	0.450***
Terrace	0.101***	0.095***	0.254***
Elevator	0.027***	0.030***	−0.014
Building number of floors	0.007**	0.006**	0.015
Building quality	0.444***	0.452***	0.311***
Presence of Caretaker	0.070***	0.067***	0.178***
Microzone	0.452***	0.458***	0.555***
Presence of garage	0.093***	0.093***	0.106**

Notes: Significance codes: ***$0 < \Pr(> |t|) < 0.001$; **$0.001 < \Pr(> |t|) < 0.01$; *$0.01 < \Pr(> |t|) < 0.05$; ****$0.05 < \Pr(> |t|) < 0.1$; $^{\dagger}0.1 < \Pr(> |t|) < 1$

Table VIII.
Characteristics linear correlations

ρ	No. of bathroom	Size	No. of rooms
Number of bathroom	1.000		
Size	0.767	1.000	
Number of rooms	0.712	0.865	1.000

Characteristic	Sample agents Estimate	Standard error	*t* value	Pr(> t)	Sample sellers Estimate	Standard error	*t* value	Pr(> t)
(Intercept)	11.972	0.099	121.083	< 2e-16***	11.424	0.427	26.752	< 2e-16***
Microzone 1	Omitted				Omitted			
Microzone 2	−0.296	0.102	−2.896	0.004**				
Microzone 3	−0.103	0.108	−0.951	0.342	0.639	0.443	1.445	0.161
Microzone 4	−0.191	0.114	−1.680	0.093****				
Microzone 5	−0.210	0.106	−1.978	0.048*	0.101	0.504	0.200	0.843
Microzone 6	−0.098	0.142	−0.690	0.490				
Microzone 7	−0.424	0.108	−3.939	0.000***	0.088	0.433	0.203	0.840
Microzone 8	−0.250	0.112	−2.242	0.025*				
Microzone 9	−0.342	0.111	−3.077	0.002**	0.142	0.480	0.295	0.770
Microzone 10	−0.545	0.112	−4.850	0.000***	−0.194	0.486	−0.399	0.693
Microzone 11	−0.503	0.105	−4.772	0.000***	0.164	0.443	0.371	0.714
Microzone 12	−0.512	0.115	−4.439	0.000***	0.458	0.438	1.046	0.305
Microzone 13	−0.234	0.158	−1.482	0.139				
Microzone 14	−0.236	0.113	−2.099	0.036*				
Microzone 15	−0.354	0.109	−3.258	0.001**	0.362	0.449	0.807	0.427
Microzone 16	0.086	0.133	0.647	0.518				
Microzone 17	−0.327	0.123	−2.657	0.008**				
Microzone 18	−0.412	0.109	−3.778	0.000***	0.169	0.527	0.321	0.751
Microzone 19	−0.776	0.103	−7.537	0.000***	−0.279	0.444	−0.627	0.536
Microzone 20	−0.735	0.112	−6.571	0.000***	−0.442	0.441	−1.002	0.326
Microzone 21	−1.034	0.102	−10.141	< 2e-16***	0.017	0.475	0.035	0.972
Microzone 22	−0.393	0.119	−3.310	0.001***	−0.166	0.438	−0.379	0.708
Microzone 23	−0.108	0.113	−0.961	0.337				
Microzone 24	−0.481	0.099	−4.854	0.000***	0.839	0.394	2.129	0.043*
Microzone 25	−0.546	0.121	−4.500	0.000***	0.100	0.484	0.206	0.838
Microzone 26	−0.669	0.107	−6.252	0.000***				
Microzone 27	−0.651	0.118	−5.531	0.000***	−0.154	0.496	−0.311	0.758
Microzone 28	−0.750	0.110	−6.830	0.000***	−0.044	0.443	−0.100	0.921
Microzone 29	−0.603	0.099	−6.084	0.000***	−0.015	0.422	−0.036	0.971
Microzone 30	−0.742	0.113	−6.561	0.000***	0.333	0.486	0.685	0.499
Microzone 31	−0.657	0.109	−6.028	0.000***	−0.014	0.418	−0.034	0.973
Microzone 32	−0.488	0.104	−4.674	0.000***	0.162	0.464	0.349	0.730
Microzone 33	−0.680	0.103	−6.624	0.000***	−0.174	0.478	−0.365	0.718
Microzone 34	−0.669	0.117	−5.696	0.000***				
Microzone 35	−0.895	0.103	−8.725	< 2e-16***	−0.388	0.423	−0.917	0.368
Microzone 36	−1.002	0.115	−8.720	< 2e-16***				
Microzone 37	−0.804	0.109	−7.409	0.000***				
Microzone 38	−0.884	0.111	−7.936	0.000***	−0.253	0.433	−0.583	0.565
Microzone 39	−0.287	0.126	−2.273	0.023*				
Microzone 40	−0.840	0.113	−7.411	0.000***				
Building quality_1	−0.101	0.059	−1.733	0.083****	−0.340	0.204	−1.667	0.108
Building quality_2	Omitted				Omitted			
Building quality_3	0.199	0.024	8.227	0.000**	−0.228	0.118	−1.922	0.066****

(*continued*)

Table IX. Regression coefficients

Characteristic	Sample agents Estimate	Standard error	*t* value	Pr(> t)	Sample sellers Estimate	Standard error	*t* value	Pr(> t)
Building quality_4	0.452	0.062	7.240	0.000***	0.132	0.343	0.386	0.703
Building quality_5	0.330	0.034	9.684	< 2e-16***	−0.025	0.182	−0.140	0.890
Size (sqm)	0.008	0.000	40.710	< 2e-16***	0.010	0.002	6.202	0.000***
Terrance_1	0.061	0.023	2.683	0.007**	−0.410	0.170	−2.416	0.023*
Garage_1	0.086	0.030	2.907	0.004**	−0.056	0.150	−0.373	0.712
Year 2011	Omitted				Omitted			
Year 2012	−0.101	0.018	−5.686	0.000***	0.055	0.105	0.521	0.607
	–				–			

Notes: Sample agents: significance codes: ***0 < Pr(> |t|) < 0.001; **0.001 < Pr(> |t|) < 0.01; *0.01 < Pr(> |t|) < 0.05; ****0.05 < Pr(> |t|) < 0.1; †0.1 < Pr(> |t|) < 1; residual standard error: 0.2892 on 1,070 df; multiple *R*-squared: 0.8637; adjusted *R*-squared: 0.8577; *F*-statistic: 144.2 on 47 and 1070 DF, *p*-value: < 2.2e-16; sample sellers: significance codes: ***0 < Pr(> |t|) < 0.001; **0.001 < Pr(> |t|) < 0.01; *0.01 < Pr(> |t|) < 0.05; ****0.05 < Pr(> |t|) < 0.1; †0.1 < Pr(> |t|) < 1; residual standard error: 0.2439 on 26 df; multiple *R*-squared: 0.9265; adjusted *R*-squared: 0.8359; *F*-statistic: 10.24 on 32 and 26 df, *p*-value: 2.393e-08

Table IX. Regression coefficients

characteristics which has a negligible explanatory power of listing prices: the unit condition, the presence of an elevator, the building number of floors, the presence of a caretaker and the presence of a garage. Despite these characteristics that are revealed to potential buyers to promote a house, they are not useful for listing prices prediction.

On the other hand, the characteristics most associated with *LP* are: size, number of rooms, number of bathrooms, location and building quality. Number of rooms, number of bathrooms and size are strongly correlated each other, as Table VIII shows. We found empirical evidence of the key role played by size, location and quality of the building when houses are put on the market; in fact, these characteristics are used for two purposes: promote the house and define the listing price. The coefficient of determination of the variable Microzone highlights the importance of Turin Microzones as real estate submarkets, as they explain 45 per cent of price variation.

Let us now compare the results on the two subsamples: *Agents* and *Sellers*. Agents showed a higher perception than sellers of the contribution of individual characteristics to selling price. In fact, in the sample *Agents*, each coefficient of determination – except the floor – is significant, whereas the characteristics' elevator, building number of floors and unit condition does not explain the listing price defined by sellers at all. Although several papers have found empirical evidence that house repair costs are usually included in the house selling price, as we mentioned in Remark 3.1, the contribution of apartment condition to listing price in negligible in the sample *Agents*, as confirmed by the coefficient of determination $R^2_{\text{Adj}} < 0.05$. This is probably due to everyday experience of agents, which sustain that buyers often decide to refurbish the apartment despite it is not necessary.

Location is an important factor for both sellers and agents. The simple regression performed with the statistical explanatory variable Microzone reveals that agents take location into account to define the listing prices, as all the Microzones coefficients are significant ($p < 0.01$). In contrast, for sellers, only the hill zone has a significant

contribution to price ($p < 0.01$). Although price statistics for each Microzone are available to sellers, sellers seem not to consider location when they put their houses on sale.

This result is confirmed by the hedonic regression analyses presented below and indicates a difference between sellers and agents awareness of the contribution of location to a house value.

The second step of work assessed the overall contribution of characteristics to listing prices, by performing two regression analyses: the first to explain agents listing prices, the second to explain sellers listing prices. The results are provided in Table IX. The set of explanatory variables used in the hedonic model includes: Microzone, building quality, size, presence of a terrace, presence of a garage and also the year of the advertisements. The list is obtained excluding from the list in Table IV the characteristics whose coefficient of determination calculated from the whole sample is $R^2_{adj} < 0.1$, i.e. floor, unit condition, elevator, building number of floor and presence of a caretaker. Indeed, we decided to include in the model also the presence of a garage, as $R^2_{adj} > 0.09$. We clearly excluded the characteristic number of rooms and number of bathrooms because they are strongly correlated with size.

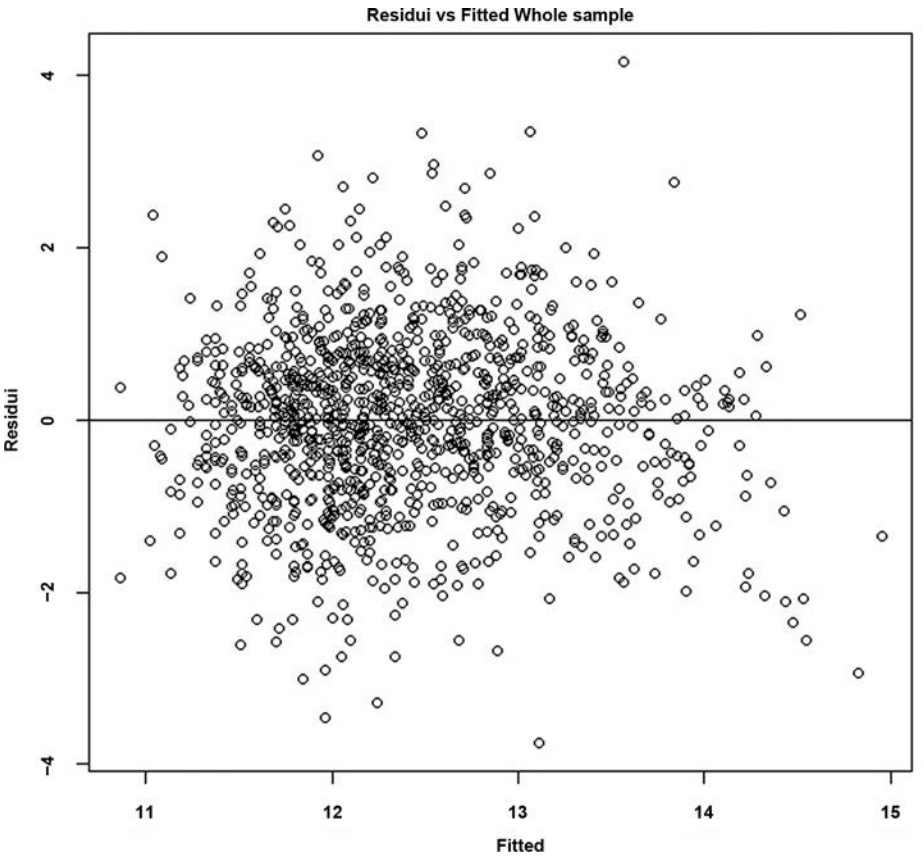

Figure 1.
Standard residuals versus fitted data for the whole sample

In both samples, we found a good fit of the hedonic model, as confirmed by the coefficients of determination, which are provided in Table IX. Nevertheless, the model failed to explain almost 15 per cent of price variation in both samples. These price variations indicate both the heterogeneity of agents and sellers reservation prices and the potential presence of unobservable factors contributing to listing price, i.e. factors not revealed in advertisements. Unobservable factors may be a source of bias in estimating the value of a house, when listing prices are used for appraisal purposes.

4.1 Remark 4.1

On the basis of the arguments discussed in Remark 3.1, we could increase the percentage of price variation explained by the hedonic model by including building condition – unobservable to buyers – in the set of regressors. Fregonara and Semeraro (2013) empirically measured the contribution of building condition to listing prices (and selling prices) in a sample of house transactions occurred in Turin in the time period of 2007-2010. They found that building condition explained almost 30 per cent of listing price variation and selling price variation. While we examined houses on sale, their analysis focus on transactions; thus, their sample does not include houses that remained

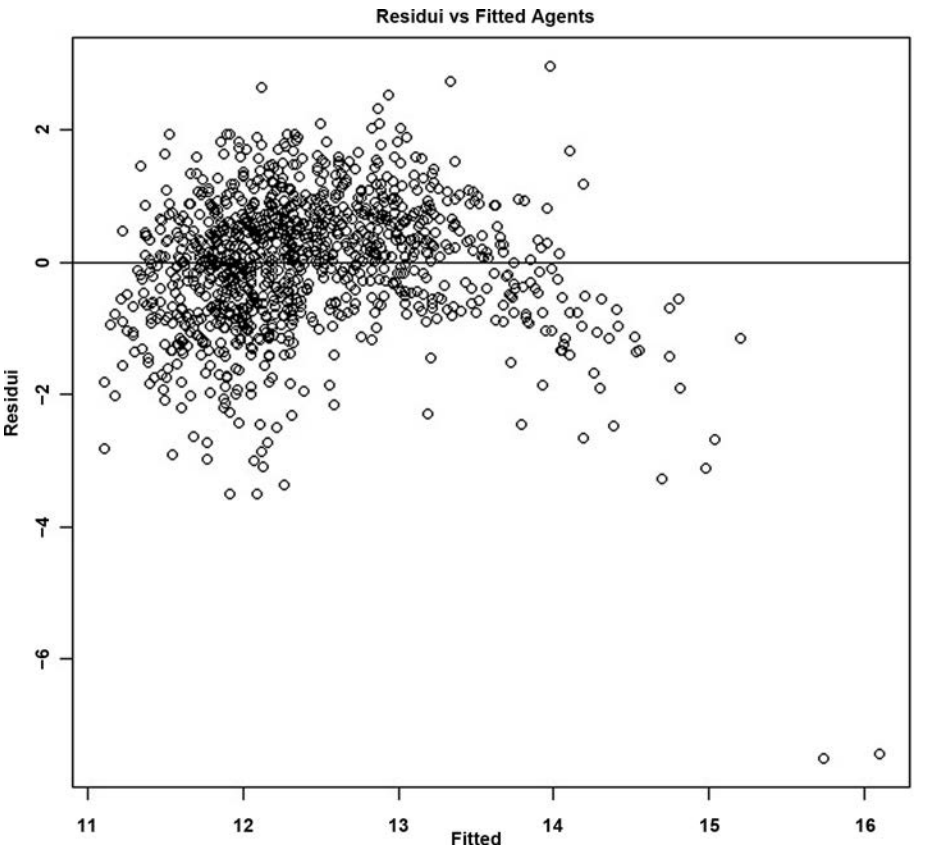

Figure 2.
Standard residuals versus fitted data for the sample *Agents*

unsold. Actually, their result exhibits the importance in a house sale of building condition, which is not revealed in advertisements.

Let us compare now the regression coefficients estimated from the sample *Agents* with the coefficient estimated from the sample *Sellers* (Table IX). We found empirical evidence that agents defined the listing price as a function of the house features. In fact, the regression coefficients of each characteristic are significant (although they are small). In contrast, sellers did not incorporate the marginal value of individual characteristics in listing price definition. Only the size, the presence of a terrace and the hill zone (Microzone 24) present a significant contribution to the listing price defined by sellers. The unawareness shown by sellers about the determinants of prices could lead the sellers to prefer to commit the sale of their house to a real estate agency, as empirically highlighted by the low percentage of houses put on sale by sellers. Notice that, the negative sign of the Microzone coefficients depends on the omitted dummy variable – Microzone 1 – that is the central city area. Microzone 16 – the central pedestrian area – shows a positive coefficient, but not significant ($p > 0.1$). Notice that agents and sellers have a different perception of the hill zone – Microzone 24 – value, whose marginal contribution to price has the different signs (both significant at level 0.05) in the two samples.

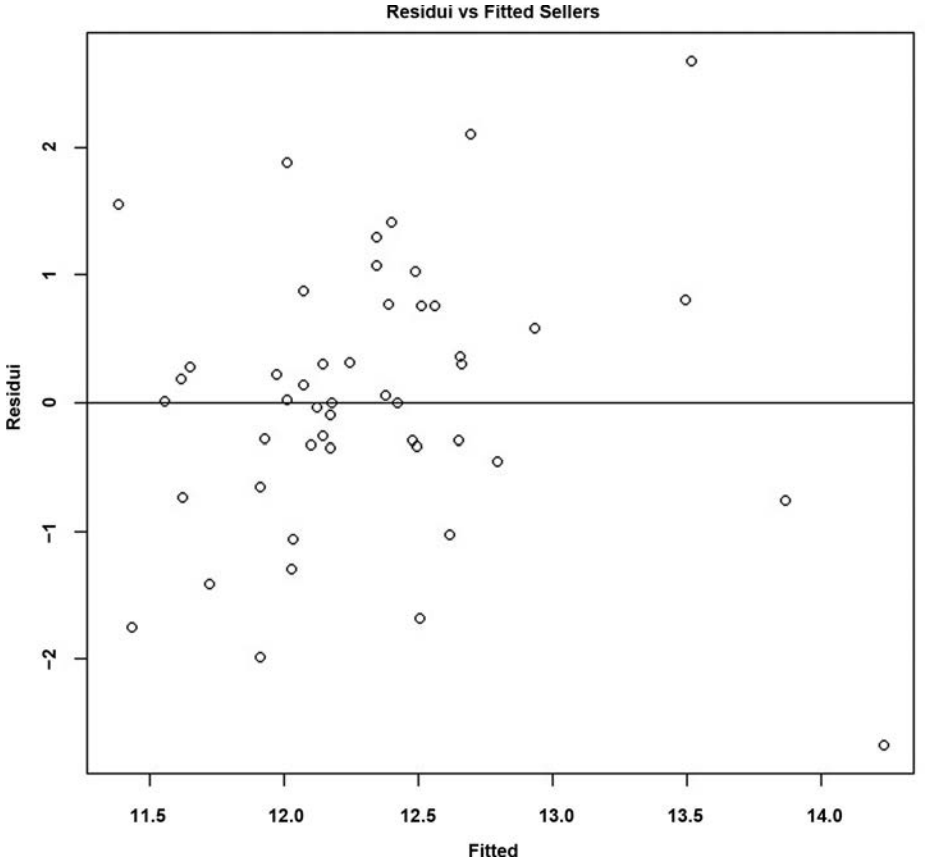

Figure 3.
Standard residuals versus fitted data for the sample *Sellers*

Also notice that low levels of building condition show a negligible contribution to price in both the samples. The presence of a garage seems to be a factor ignored by sellers. Lastly, we notice that listing prices in the sample *Agents* decreased in 2012, according agents' perception of a negative trend of prices in the past years. Sellers seem to be unaware of the prices trend.

We end this section with the analysis of residuals of the three regressions performed. We tested the residual means from the three samples and the Student's test did not reject the null hypothesis, i.e. the mean of residuals is zero, with p-values $p = 1$.

Shapiro-Wilk test rejected normality of residuals from the whole sample and from sample *Agents* with $p < 0.01$. The p-value calculated from the smallest sample *Sellers* is $p > 0.01$, the test rejected normality at level 0.05 but it did not reject normality at level 0.01. Furthermore, Figure 3 shows that the sample *Sellers* does not present any pattern, as well as the whole sample. Thus, the sample *Sellers,* shows the smallest deviation from normality, enforcing the significance of the results discussed above.

The plots of standard residuals versus fitted data and the Q-Q plots (Figures 1-6) show some patterns for the sample *Agents*. These results indicate that also the residual distributions reflect heterogeneity of agents and sellers reservation prices, as evidenced by the tail of their distributions.

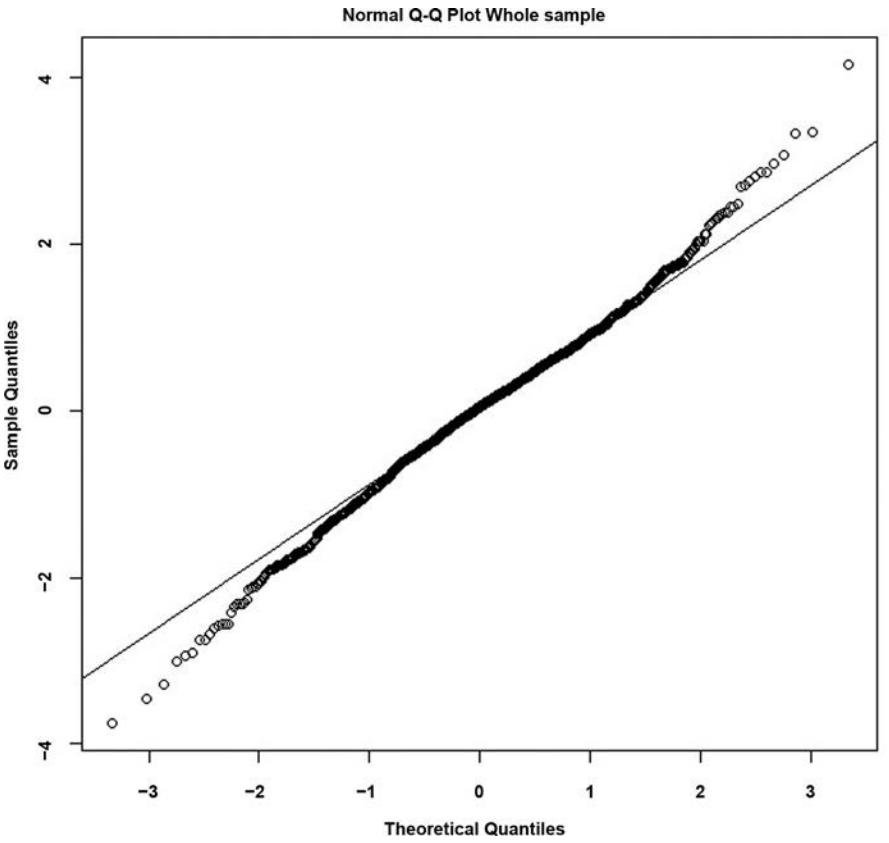

Figure 4.
Q-Q plots for the whole sample

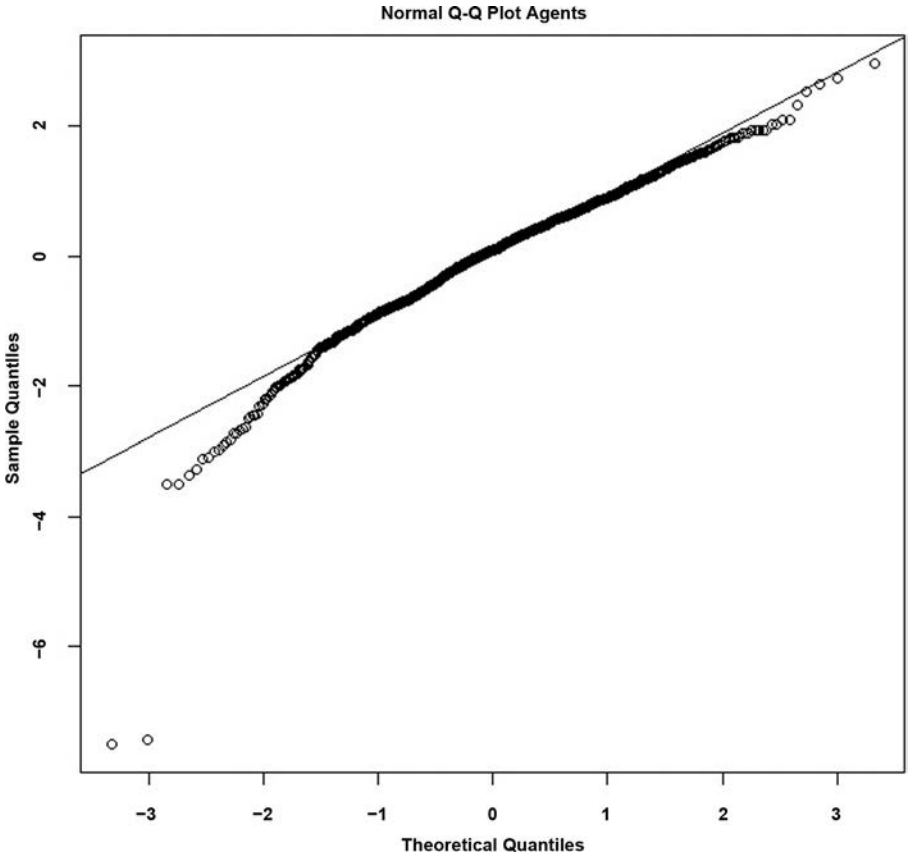

Figure 5.
Q-Q plots for the sample *Agents*

Conclusion

In the Italian real estate market, listing prices are extremely important, as selling prices are not public information. As a consequence, researchers, appraisers and real estate companies use listing prices for studying the market and estimating the value of houses. We analyzed listing prices and their determinants to address their employment in real estate valuations. Real estate advertisements include listing price, location and a list of the houses features, which we name observable characteristics.

We performed a hedonic empirical analysis in two steps. First, we measured the contribution to price of individual observable characteristics. Second, we performed a traditional hedonic regression analysis to assess the overall contribution to listing prices of information in advertisements.

We considered both houses put on sale by agents and sellers. Both samples revealed the potential presence of unobservable factors that affect listing prices; in fact the model explains almost 83 per cent of listing price variation in both cases. This fact outline that the information in advertisements could be not sufficient for accurate prices predictions. The regression analyses also revealed the presence – in both cases – of observable characteristics, whose marginal contribution to price is negligible (e.g. the presence of a

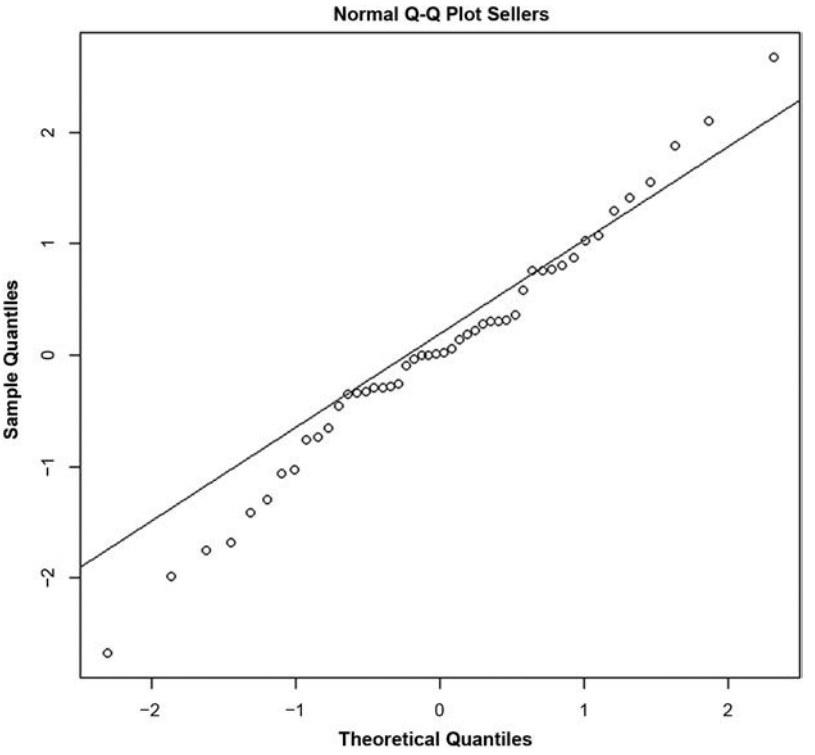

Figure 6.
Q-Q plots for the sample *Sellers*

caretaker). Agents and sellers showed a different understanding of individual house characteristics contribution to price. The regression coefficients of the sample agents are all significant ($p < 0.01$). Therefore, the observable characteristics building quality, location, size, presence of a terrace, presence of a garage and year of the advertisements are used by agents to define the listing price, according to the hedonic hypothesis. In contrast, the regression coefficients estimated on the sample *Sellers* pointed out that the sole characteristic with a high level of significance is size. In that, we found empirical evidence of the sellers' unawareness of the importance of the house features contribution to prices, as they only use the characteristics to describe the property in their web advertisements. These results indicate that the predicting power of house features is higher in houses put up by real estate agents than in houses put up by sellers.

Note

1. Presidential Decree 138/1998 and ensuing Regulation issued by the Ministry of Finance.

References

Anglin, P.M., Rutherford, R. and Springer, T.M. (2003), "The trade-off between the selling prices of residential properties and time-on-the-market: the impact of price setting", *Journal of Real Estate Finance and Economics*, Vol. 26 No. 1, pp. 95-111.

Beracha, E. and Seiler, M.J. (2013), "The effect of listing strategy on transaction selling prices", *Journal of Real Estate Finance and Economics*, Vol. 49 No. 2, pp. 237-255.

Bourassa, S.C., Cantoni, E. and Hoesli, M. (2007), "Spatial dependence, housing submarkets, and house prices", *Journal of Real Estate Finance and Economics*, Vol. 35 No. 1, pp. 143-160.

Bourassa, S.C., Cantoni, E. and Hoesli, M. (2008), "Predicting house prices with spatial dependence: impacts of alternative submarket definitions", Swiss Finance Institute Research Paper No. 08-01, available at: http://ssrn.com/abstract=1090147 or http://dx.doi.org/10.2139/ssrn.1090147

Bourassa, S.C., Hoesli, M. and Peng, V.S. (2003), "Do housing submarkets really matter?", *Journal of Housing Economics*, Vol. 12 No. 1, pp. 12-28.

Box, G.E.C. and Cox, D.R. (1964), "An analysis of transformations", *Journal of the Royal Statistical Society. Series B. (Methodological)*, Vol. 26 No. 2, pp. 211-252.

De Wit, E.R. and Van der Klaauw, B. (2013), "Asymmetric information and list-price reductions in the housing market", *Regional Science and Urban Economics*, Vol. 43 No. 3, pp. 507-520.

Fregonara, E., Semeraro, P. (2013), "The impact of house characteristics on the bargaining outcome", *Journal of European Real Estate Research*, Vol. 6 No. 3, pp. 262-278.

Gini, C. (1912), *Variabilità e mutabilità*, C. Cuppini, Bologna.

Glower, M., Haurin, D.R. and Handershott, P.H. (1996), "Selling time and selling price: the influence of seller motivation", *Real Estate Economics*, Vol. 26 No. 4, pp. 719-740.

Goodman, A.C. and Thibodeau, T.G. (1998), "Housing market segmentation", *Journal of Housing Economics*, Vol. 7 No. 1, pp. 121-143.

Horowitz, J.L. (1992), "The role of the list price in housing markets: theory and an econometric model", *Journal Of Applied Econometrics*, Vol. 7 No. 2, pp. 115-129.

Kauko, T.J. (2006), "Urban housing patterns in a tide of change", *Spatial Structure and Residential Property Values in Budapest in a Comparative Perspective*, Amsterdam IOS Press, Delft University press, Vol. 8.

Knight, J.R. (2002), "Listing price, time on market, and ultimate selling price: causes and effects of listing price changes", *Real Estate Economics*, Vol. 30 No. 2, pp. 213-237.

Knight, J.R., Miceli, T. and Sirmans, C.F. (2000), "Repair expenses, selling contracts and house prices", *Journal of Real Estate Research*, Vol. 20 No. 3, pp. 324-336.

Knight, J.R. and Sirmans, C.F. (1996), "Depreciation, maintenance, and housing prices", *Journal of Housing Economics*, Vol. 5 No. 4, pp. 369-389.

Knight, J.R., Sirmans, C.F. and Turnbull, J.K. (1998), "List price information in residential appraisal and underwriting", *Journal of Real Estate Research*, Vol. 15 No. 1, pp. 59-76.

Lerman, R.I. and Yitzhaki, S. (1984), "A note on the calculation and interpretation of the Gini coefficient", *Economic Letters*, Vol. 15 Nos 3/4, pp. 363-368.

Pace, R.K., Barry, R. and Sirmans, C.F. (1998), "Spatial statistics and real estate", *The Journal of Real Estate Finance and Economics*, Vol. 17 No. 1, pp. 5-13.

Robertson, K. and Doig, A. (2010), "An empirical investigation of variations in real-estate marketing language over a market cycle", *Housing, Theory and Society*, Vol. 27 No. 2, pp. 178-189.

Rodriguez, G. and Siret, D. (2009), "The future of houses: what real-estate ads tell about the evolution of single-family dwellings", *International Journal of Architectural Research*, Vol. 3 No. 1, pp. 92-100.

Rosen, S. (1974), "Hedonic prices and implicit markets: product differentiation in pure competition", *Journal of Political Economy*, Vol. 82 No. 1.

Sirmans, G.S., Macpherson, D.A. and Zietz, E.N. (2005), "The composition of hedonic price models", *Journal of Real Estate Literature*, Vol. 13 No. 1, pp. 1-44.

Song, S. (1995), "Determinants of bargaining outcome in single-family housing transaction: an empirical examination", *Urban Studies*, Vol. 32 No. 3, pp. 605-614.

Yavas, A. and Yang, S. (1995), "The strategic role of listing price in marketing real estate: theory and evidence", *Real Estate Economics*, Vol. 23 No. 3, pp. 347-368.

Appendix

A.1 Gini concentration index

We recall here the normalized Gini concentration index (Gini, 1912; see also Lerman and Yitzhaki, 1984), that is defined as follows:

$$G = \frac{n-1}{n}\left(1 - \sum_i f_i^2\right),$$

where f_i is the relative frequency of the i*th* level, n is the number of levels. The coefficient G takes values in the interval [0,1], where $G = 0$ indicates minimal heterogeneity (all the observations belongs to the same level) and $G = 1$ Indicates maximal heterogeneity (the observation are equally distributed among levels).

A.2 Box-Cox transformation

The Box-Cox transformation (Box and Cox (1964)), to recover normality of the transformed variable, is the following parametrical family of transformations of the dependent variable:

$$Y^{\gamma} = \begin{cases} \dfrac{Y^{\gamma} - 1}{\gamma}, & \gamma \neq 0 \\ \log Y & \gamma = 0 \end{cases}$$

Notice that the logarithmic transformation belongs to the family of Box-Cox transformations, corresponding to the parameter specification $\gamma = 0$.

About the authors

Rocco Curto, Director of the Department of Architecture and Design, Turin Polytechnic. Since 1999, Full Professor in Property Valuation, Turin Polytechnic. Dean of the II Faculty of Architecture, Turin Polytechnic (2006-2012). Member of the management committee of the Territorial Agency, Italian Ministry of Economy and Finance (2007-2008). Director of the Master's Course in Real Estate: Territorial Planning and Property Market c/o COREP and Turin Polytechnic (1999-2012). Since 2004, Director of the Master's Course in Management of Cultural and Environmental Heritage c/o COREP and the Turin Polytechnic. Since 1999, Head of the Turin Real Estate Observatory. Consultant/Advisor for Public Authorities, often as Director. Research Topics: property market; application of statistical techniques (forecasting and probability), mass appraisals; functional and territorial segmentation of the property market; economic-financial evaluation of projects and plans; enhancement and management of historical and architectural heritage; analysis of real estate investments and risk management; reform of the Land Register and revision of the cadastral rent.

Elena Fregonara, Degree in Architecture, PhD, in Urban Planning and the Real Estate Market. Associate Professor in Real Estate Appraisal and Economic Evaluation of Projects since 2006, at Architecture and Design Department, Politecnico di Torino. She is Vice-responsible of Turin Real Estate Market Observatory. Lecturer of many teachings, among others: Appraisal and Professional Practice, Economic Evaluation of Projects and Plans. Member of the scientific committee at Master in "Urban Planning and The Real Estate Market" and at Master in "Management of Cultural and Environmental Heritage". Her research activity has been focussed on: economic evaluations of assets and projects, in private and public context; economic feasibility of projects and of capital investment in the real estate sector under risk and uncertainty; the real

estate market analysis and monitoring of values and dynamics (statistical and econometric models); the Cadastral review, in particular about the definition of equal microzones for Turin municipality.

Patrizia Semeraro is the Assistant Professor of Real Estate Appraisal and Economic Evaluation of Projects since 2011 at Architecture and Design Department, Politecnico di Torino. She is a graduate in Mathematics of University of Turin and holds a PhD in Mathematics at University of Turin. She currently teaches *Real Estate valuations: theory and Methods* at Politecnico di Torino, *Market Segment*ation at a graduate Master "Management of Cultural and Environmental Heritage" and *Mass Appraisal* at a graduate Master "Urban Planning and The Real Estate Market" of Politecnico di Torino. She taught Statistics at a graduate Master in Economics, Coripe Piemonte-Collegio Carlo Alberto. Her research activity has been focussed on: applied probability, Lévy processes for finance and recently real estate markets. She has published in international finance and applied mathematics Journals including *Journal of Applied probability*, *Journal of Theoretical and Applied Finance*, *Mathematics of Operations Research*, *Quantitative Finance*, *Journal of Computational and Applied Mathematics* and *Journal of European real Estate research*. Patrizia Semeraro is the corresponding author and can be contacted at: patrizia.semeraro@polito.it