uc3m | Universidad Carlos III de Madrid

University Degree in ECONOMICS

Academic Year (2018-2022)

Bachelor Thesis

"Hedonic price models for rental housing in Madrid City vs South suburbs"

Author: Celia Moreno González

Advisor: Dr. Carlos San Juan Mesonado

Getafe (Madrid), June 2022

cc BY NC ND

This work is licensed under Creative Commons **Attribution – Non Commercial – Non Derivatives**

INDEX

FIGURES INDEX

TABLES INDEX

ABSTRACT:

One of the characteristics of the housing market is heterogeneity, in order to know which features most affect the rental price, an appropriate research must be developed.
Madrid is considered to be a very demanded, polycentric city for being the capital of Spain. This is why an analysis of the main characteristics that fix the rental price I the areas of Madrid City and South Madrid has been carried out using a Hedonic price model. For it, a total amount of 3.069 cross-sectional data from early 2022 has been collected. In the estimations OLS, 2SLS and GMM have been used to correct endogeneity problems. Landlords and tenants can use this approach in order to establish the rental price of their homes based on the variables that most affect them.
According to the existing literature, there are no previous investigations covering the very actual period of data collection and areas treated in this analysis.
Results of the models show that the same house located in the South is consistently cheaper than in the City. The university core has a greater increase in the price when having one more bathroom or pool in the dwelling compared to the non-university areas.

Key Words: Rental Price, Housing, Hedonic price model, Madrid, University area.

1. INTRODUCTION

"***Heterogeneity**, **durability** and **spatial fixity***" *(Tung-Leong, K W CHAU*, 2003),* these are some of the characteristics that describe the housing market, a market which can be considered complex in such a way that generalization between regions and time is not feasible.

Dwellings are assets that constitute a large investment in the life of a consumer, either for the use of the investor or as speculative instruments in the form of rental to third parties. In this case, landlords make the investment decision expecting to receive periodical revenues. To maximize these benefits, it is essential to know which are the characteristics determining the rental price.

The rental property price is not only affected by intrinsic attributes in each home but also by other external factors such as the environment or the location. One of the main characteristics determining the final rental price is the area where the property is located, especially in big cities like Madrid, London or Berlin where prices vary among different areas. Previous research has focused its analysis on high populated areas of Turkey *(Sibel, 2008*), China *(Hanink, Cromley & Ebenstein, 2010),* Ecuador (*Zambrano-Monserrate, 2015*), Vietnam (*NGUYEN, Quoc PHAN, Tri Van TRAN, Kiem Viet TRAN, 2020*), Colorado (*Yazdani, 2021*) or Moroco (*Firano, A. Filiali, 2021*).

These studies use the *Hedonic Price Models* to perform their analysis. Hedonic Price Models estimate the monetary value of a good through the sum of its most important characteristics. Historically, they were introduced into the housing market firstly by *(Lancaster, 1966)* and more deeply by *(Rosen, 1974)*. In this market, they allow estimating the impact of the characteristics of the house in the final rental price. Besides the location, other relevant variables are square meters, rooms, bathrooms and other dummy variables such as a terrace or air conditioner.

This analysis is important to **policymakers**, **suppliers**, as well as for **demanders.** On the supplier side, this study helps to answer questions such as: Is it more profitable to buy a house in a University municipality? How much is the rental price likely to increase if an air conditioner is installed in the house? From the demander point of view, it provides a piece of empirical evidence about the expected rental price according to the particular attributes of each home. Housing markets are commonly regularized by the public institutions, so this study is also relevant to policymakers in order to better understand and therefore regulate the market.

This project analyzes the **rental price** in the areas of **City** and **South Madrid.** According to previous literature, these areas have not been analyzed yet for such an actual period.

The **methodology** used in this econometric study is a multiple linear regression, calculated using the Ordinary least squares (**OLS**), Two Stage Least Squares (**2SLS**) and Generalized Methods of Moments (**GMM**). The steps from *(Wooldridge, 2008)* have been followed during the econometric analysis. For the evaluation and models, actual market data has been considered, collecting more than 3000 entries from the main housing supply platform in Spain "*El Idealista*[1]".

Given the heterogeneity of Madrid´s rental housing market, which has a wide range of supplies, commodities, prices and locations, a specific study must be done. The **main target** of this work is to find which characteristics mostly affect the rental prices, which characteristics are not significant and what is the difference between living in City Center or South suburbs of Madrid.

2. OBJECTIVES AND STRUCTURE:

The main objective of this final bachelor's project is to identify the most significant characteristics that fix the rental price in the areas of Madrid City and South Madrid by applying a Hedonic price model.

In order to reach the goal, we have identified and achieved the following sub-objectives.

1. Study of the state of art focusing on previous research works and literature.
2. Analysis of the rental market in Madrid, evaluation and selection of the areas to cover and identification of parameters to take into account.
3. Data collection from "*El Idealista*" portal and creation of the database.
4. Review of acquired data.
5. Elaboration of the hedonic price models by creating the regressions script in *Stata* for the selected areas.
6. Evaluation of results and conclusions.
7. Preparation of the final bachelor project memory.

The structure works as follows: Section 3 is dedicated to the literature review, section 4 explains the areas of interest as well as descriptive statistics of the data. Section 5 focuses on the data used, section 6 analyzes the different econometric models and the methodology used, section 7 is dedicated to the results and finally, section 8 summarizes the most important facts of the research.

---

[1] "*El Idealista*", is a Spanish Company with more than 20 years of experience as an intermediary between buyers and sellers of housing purchases and rentals.

3. LITERATURE REVIEW

There are numerous investigations that focus their attention on explaining and predicting the behaviour of rental prices. Because of the heterogeneity of this market, each location needs to be treated in a particular way. This section details the studies that have served as the basis for carrying out this final degree project, as well as the variables to be included in the model, and in the general understanding of the subject.

As starting a point, it is not clear which is the beginning of hedonic price models but it must be said that these models can be examined in the housing market thanks to the "*Consumer Theory*" (*Lancaster*, *1966*) and later the model of (*Rosen*, *1974)*.

The literature review of *(Tung-Leong, K W CHAU*, 2003)* emphasized the difficulty on making this type of analysis as they are "*Constrained by the availability of data*" and describes hedonic price models as the "*major scientific method"* to analyze the effect of each characteristic on the price. After reviewing several econometric models the author created Table 1, this table serves as a starting point to know which are the most representative characteristics of a home. In it, the signs of the different variables that affect the majority of the analyzed locations have been included. Thanks to this, today we can know with greater certainty what are those values that mark the general rental prices.

Table 1: List of Commonly used housing attributes in Hedonic Price Models

Attribute		Expected effect on housing price
Locational	Distance from CBD	-ve
	View of the sea, lakes or rivers	+ve
	View of hills/valley/golf course	+ve
	Obstructed view	-ve
	Length of land lease	+ve
Structural	Number of rooms, bedrooms, bathrooms	+ve
	Floor area	+ve
	Basement, garage, and patio	+ve
	Building services (e.g. lift, air conditional system etc)	+ve
	Floor level (multi-storey buildings only)	+ve
	Structural quality (e.g., design, materials, fixtures)	+ve
	Facilities (e.g., swimming pool, gymnasium, tennis court)	+ve
	Age of the building	-ve
Neighbourhood	Income of residents	+ve
	Proximity to good schools	+ve
	Proximity to Hospitals	?
	Proximity to Places of worship (e.g., mosques, churches, temples)	+ve
	Crime rate	-ve
	Traffic/airport noise	-ve
	Proximity to Shopping centers	?
	Proximity to Forest	?
	Environmental quality (e.g., landscape, garden, playground)	+ve

+ve – positive impact oh housing prices; -ve – negative impact oh housing prices
? – varies from place to place, the actual effect is an empirical question

SOURCE: (Tung-Leong, K W CHAU, 2003) p.13*

"*Homes are heterogeneous goods, the environment and the consumer vary in each area*" this text extracted from (*Stacy, Macpherson & Zietz, 2005)* leads us to think about the importance on particularizing the models for each location. Even though there are several characteristics which are normally important to explain most of the rental prices (see previous paragraph), there are some peculiarities of each area given the cultural and geographical differences. For example: several studies in Latin American add characteristics such as water supply (*Zambrano-Monserrate, 2015*), access to hot water or cable TV (*Sibel, 2008*) due to the different needs in comparison which more developed areas where these commodities might not be considered.

The authors (*Dimitra Kavarnou and Anupam Nanda*) joint several times to analyze two different areas. The first one The Greek Islands *(Dimitra Kavarnou, Anupam Nanda, 2014)*. This study is very unique because it was a newly-studied area. The results follow the common literature, for example, the living space and land area showed a positive impact on the final price. Peculiar variables that were included in this model are "Age", "Presence_Hospital", "Tourism" or "Presence_Airport". During their second hedonic price analysis (*Dimitra Kavarnou, Anupam Nanda, 2015*) Panama´s housing market was investigated. They divided the country into 11 zones, and quarterly data was involved from 2005 to 2014. A positive and highly statistically significant coefficient for the variables "Bathrooms", "Ln(sqm)", and "Service_room" were shown during the estimation whereas "No covered parking" showed a negative coefficient.

Over the years, the needs of the population changed and therefore more factors affect the rental prices. *(Stacy, Macpherson & Zietz, 2005)* paper investigates how the time that the home is on the market affects its price; this relation seems to be negative, that is, the longer it is in the market, the lower the sale price.

Regarding the appraisal of goods, despite the fact that this work focuses on renting them, it is important to mention this paper by (*Kirill, Pröllochs, 2021).* These authors analyzed variables such as "Floor plans" "relative size" and "positioning between rooms" in their hedonic price model. This leads us to think about what other variables could be relevant, new and valued when renting or buying a house which are not so intuitive to explain housing prices.

(Xiaochen, Chulasai, Phuangsaichai, 2011) collected a sample of data for 2009 in the Chinese city of Kunming. Their conclusion was the high importance of structural and locational attributes. Being a city with a high population density and with a business nucleus, variables such as the distance to transport or the CBD are highly relevant.

Taking as a reference the lines of research that have already been followed, this work contributes to the existing literature by analyzing, with current data (February-April 2022) the rental market in Madrid-City and South-Madrid, through a hedonic Price model using the appropriate characteristics following the necessities of this area.

4. AREAS AND DESCRIPTIVE STATISTICS:

a. AREAS:

Madrid´s community is made up of 179 municipalities and can be divided into 5 areas. North-Madrid, Madrid-City, South-Madrid, East-Madrid and West-Madrid *(Comunidad de Madrid, 2009)*. The areas to be threated during the project are **Madrid-City** and **South-Madrid.**

Inside Madrid-City 21 districts are found, for this final bachelor project the districts of **Central Madrid** and **Chamberí** have been selected.

Central-Madrid is constituted by 6 neighborhoods: Universidad, Justicia, Sol, Cortes, Embajadores and Palacio whereas Chamberí is formed by another 6 neighborhoods: Gaztambide, Arapiles, Trafalgar, Almagro, Rios Rosas and Vallehermoso.

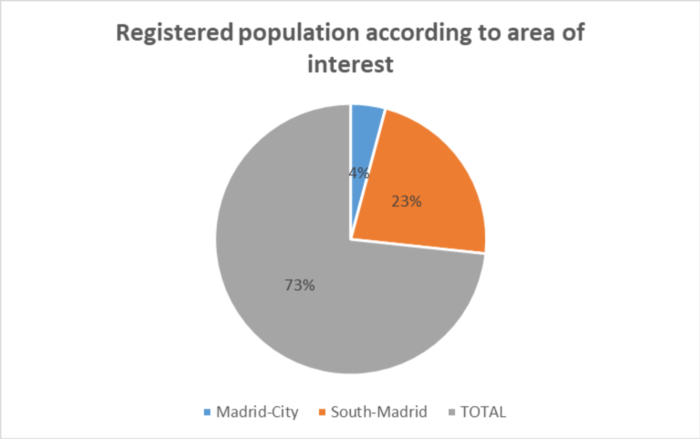
South-Madrid, it is composed by 39 municipalities but the only ones with supply of rental houses and therefore the ones relevant for this project are: Alcorcón, Leganés, Getafe, Móstoles, Fuenlabrada, Pinto, Arroyomolinos, Parla, Valdemoro, Ciempozuelos, Batres, Griñón, Torrejón de Velasco and Humanes.

Prior to the model and analysis of the database, an analysis of the size of the zones in terms of their population will be done taking as a reference the number of registered residents in each area. The selection of registered residents is taken as reference given that the Madrid city council states "*Anyone who lives in Spain is **obliged to register in the municipal register where they habitually reside***". Therefore, it is understood, that the number of registered people represents the number of people commonly living in each neighborhood.

As seen in Figure 1, the **total registered population** in the Community of Madrid is 6.779.888 citizens. Madrid-City constitutes a 4% of the total with 281.358 people whereas South-Madrid represents 23% of the total population with 1.527.838 *(Ayuntamiento de Madrid, 2022)*

In order to be more detailed within the neighborhoods a deeper analysis has been made. In one hand, inside the area of Madrid-City we find that the neighborhoods with more registered people are Embajadores (33% of the total) and Río Rosas (20% of the total). In the other hand, in the area of South-Madrid the municipality with more registered people is Mostoles (13% of the total). In order to see this more graphically see APPENDIX A.

Figure 1: Registered population according to the areas of interest



Own elaboration based on (Ayuntamiento de Madrid, 2022)

b. DESCRIPTIVE STATISTICS OF THE DATABASE

The database has been created from the advertisements on the **reference web portal "El Idealista**" during the period between **February to April 2022**. The data has been extracted manually and has been dumped into Excel, collecting the information of the variables shown in Table 2.

Why has "El Idealista" been selected and not another real estate web portal? It is the leading web portal for purchase and rental offers in Spain. According to *(Similar web, 2022)* it leads the ranking of "*most visited Business and Consumer Services Websites Ranking*"

Data has been collected for a total of **3.069 rental houses,** specifically 2.536 in Madrid-City and 533 in South-Madrid. These numbers constitute the totality of the rental homes offered in the selected areas on the mentioned web page during the period of collection.

Table 2: List of variables collected, description and type

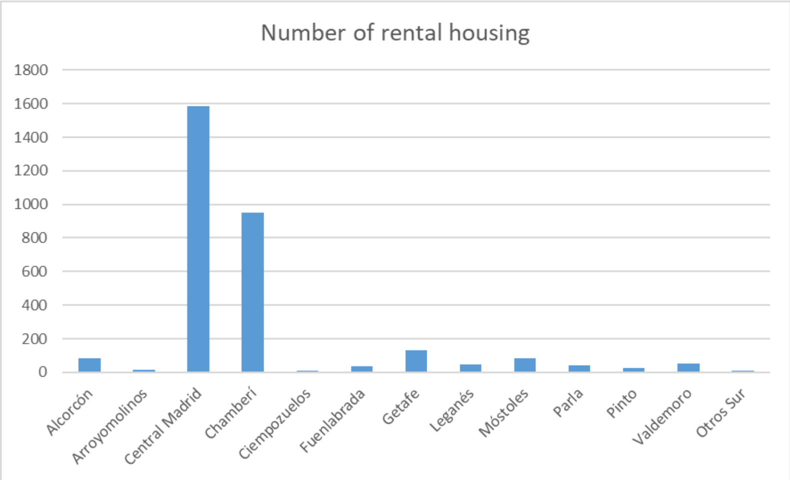
VARIABLE	DESCRIPTION	TYPE
Price	Total rental price	Cuantitative variable, continuous.
District	District	Cualitative variable, dummy.
SQM	Square Meters	Cuantitative variable, continuous
Rooms	Number of Rooms	Cuantitative variable, discrete
Bathrooms	Number of Bathrooms	Cuantitative variable, discrete
Terrace	Terrace	Cualitative variable, dummy
A/C	Air Conditioner	Cualitative variable, dummy
Garage	Garage	Cualitative variable, dummy
Pool	Pool	Cualitative variable, dummy
Elevator	Elevator	Cualitative variable, dummy

Furnished[2]	Furnished	Cualitative variable, dummy
Location	Interior or Exterior	Cualitative variable, dummy
Garden	Garden	Cualitative variable, dummy
House/Terraced house/Apartment	House (detached and semi-detached), Terraced House (duplex) or apartment	Cualitative variable, dummy
Floor	Number of floor the house is located in	Cuantitative variable, discrete

Own elaboration

After introducing the data, a prior approach of the data collected and subsequently used in the model will be carried out.

Figure 2: Number of rental housing



Own Elaboration

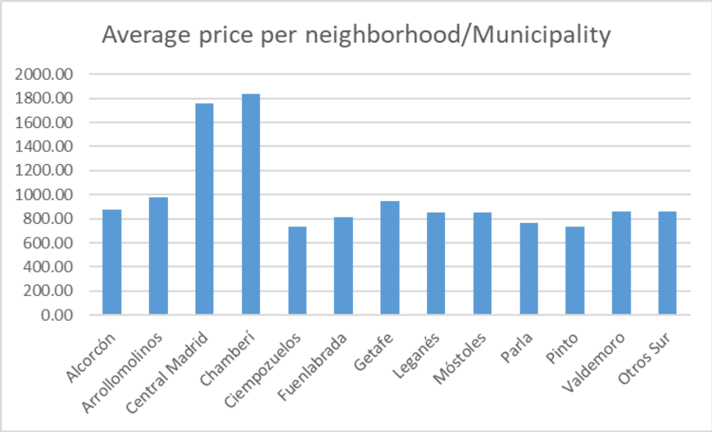
Figure 2 shows that the areas with the most offers of rental houses are Central Madrid and Chamberí. Therefore, we can conclude that the City offers much more renting houses compared to the South.

Talking about the **average rental price** of houses, Figure 3 shows the average rental price per municipalities.

Madrid-City has an average of 1.761 Euros in the Cental district and 1.840 Euros in Chamberí whereas the South gets a value of 866 Euros which is less than a half of the prices in Madrid-City. Inside the South, the Municipality with lower supply price is Pinto with an average of 734.92 Euros followed by Ciempozuelos.

[2] Initially, the variable "Furnished" was collected as a nominal qualitative variable which takes a value of 0 when it is complete unfurnished, 1 when it is completely furnished, and 2 when it has an equipped kitchen. As discussed in the model section, this variable was converted into a dummy variable before including it in the model to avoid suffering from multicolinearity.

Figure 3: Average price per neighborhood/municipality

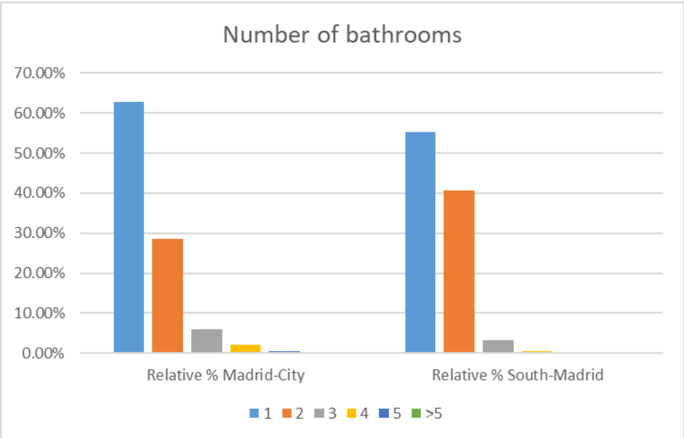


Own Elaboration

Inside both of the districts of Madrid-City, the neighborhood with higher average supply price of rental housing are Justicia in the Center and Almagro in Chamberí. whereas, the neighbors with a lower average price are Universidad in the Center and Arapiles in Chamberí.

Regarding other characteristics such as the "Number of bathrooms" and the "Number of Rooms", the following is observed.

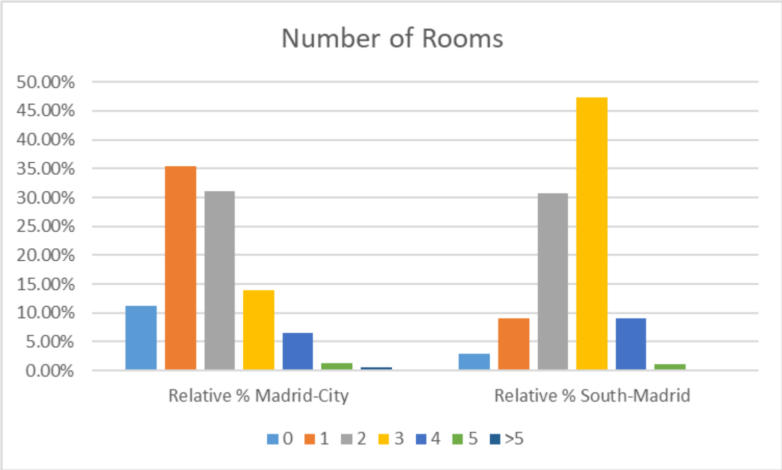
Figure 4: Number of bathrooms



Own Elaboration

Figure 4 shows the **majority of homes** offered in both areas own **one or two bathrooms** having a smaller differential between 1 and 2 in the South than in the City. Regarding the room trend shown in Figure 5, many differences can be inferred. Most of the supply of rental houses in the City have 1 or 2 rooms whereas in the South the mode is having 2 or 3 rooms. From this, we can intuit houses in the **South tend to have more rooms than the ones in the City.** It is important to look at the houses with "0" rooms, we understand these houses as "studio style" which normally have few square meters. More than 10% of the houses in the City area are studios, which compared with the percentage of this same houses in the South (less than 5%) emphasize the idea about houses in the City having less rooms.

Figure 5: Number of Rooms

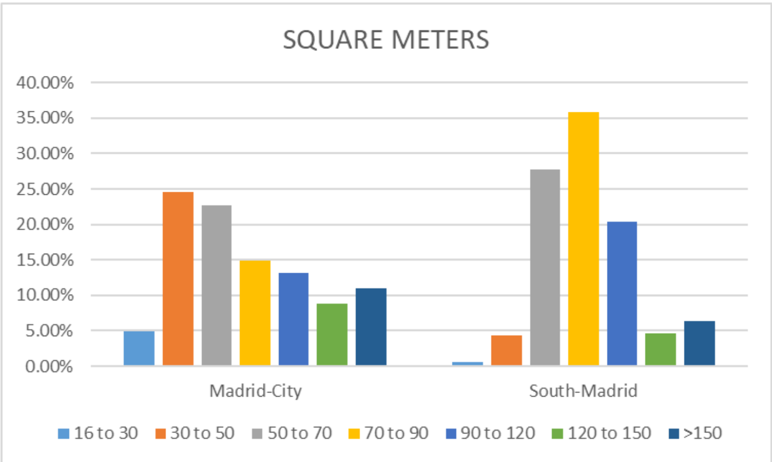


Own Elaboration

These two variables explained before can show more relevancy by looking at "Squared meters":

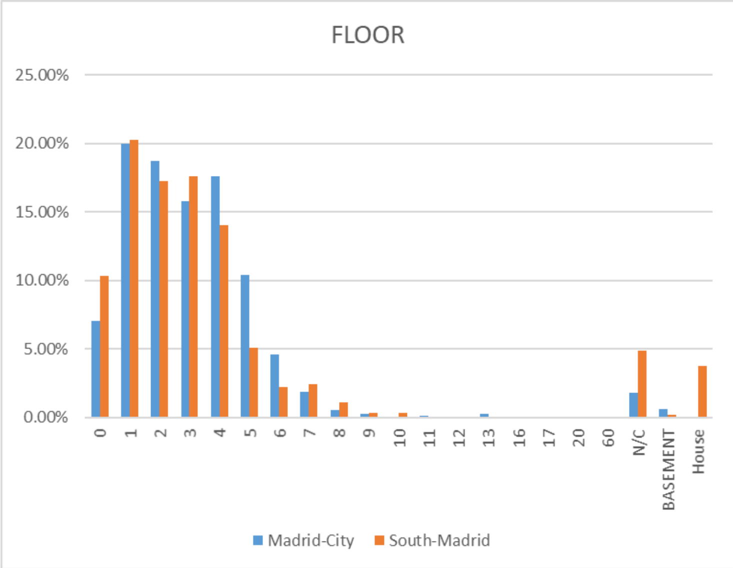
Supplied rental houses in the City commonly have between 30 and 50 square meters while the ones in the South tend to have between 70 to 90 square meters, as it is illustrated in Figure 6. Looking at some of the relations we can highlight the following: there are very few houses in the South with 16 to 30 sqm whereas in the City there is almost 5% of them which have this size. Moreover, over 20% of the houses in the South have between 90 to 120 sqm whereas in the City this percentage is lower than a 15%. This leads to conclude that houses for rent in the **South** tend to have **relatively more square meters than** the ones in the **City**.

Figure 6: Square meters



Own Elaboration

Figure 7: Floor



Own Elaboration

As shown in Figure 7**,** most of the supplied rental houses are between the **first and seventh floors**. Houses/Flats are only available for rent in the South whereas there are several Terraced House supplied in both.

Talking about the remaining variables to be observed in the subsequent models, several data should be highlighted:

It is observed that, in the City, 69% of the dwellings have **air condinitioner** whereas in the South less than 50% have it. The availibility of **garage** when renting a house is very rare in the City while in the south nearly 50% of the rental houses supplied have garage.

Having **pool** and **garden** is more common in the South than in the City being nearly zero in the City for both features. A common feature presented in more than 70% in both areas is Elevator.

Furnished rental houses are more present in the City than in the South were less than a 9% are fully furnished. The most common way of founding the rental houses are partially furnished meaning having the kitchen installed but not the rest of the house.

The **location** of the houses are also something interesting to look at, both areas have more houses offered with exterior orientation, eventhough this holds for both there are relatively more “Interior” houses in the City than in the South.

There are less houses with **terrace** in the City than in the South, where the trend is to have this commodity.

Finally, analyzing the type of housing we can see **Houses** [3]are only offered in the South, and **Terraced Houses** only consitute 1% of the houses for rent in the City and 2% in the South. We can conclude the trend in both areas is to live in an **apartment**.

Table 3: Relative frequency of the dummy variables

AIR CONDITIONER	CITY	RELATIVE FREQUENCY	SOUTH	RELATIVE FREQUENCY
Present	1749	69%	247	46%
Not present	787	31%	286	54%
GARAGE				
Present	177	7%	243	46%
Not present	2359	93%	290	54%
POOL				
Present	114	4%	158	30%
Not present	2422	96%	375	70%
LOCATION				
Interior	701	28%	26	5%
Exterior	1737	68%	506	95%
N/C	98	4%	1	0%
GARDEN				
Present	90	4%	191	36%
Not present	2446	96%	342	64%
ELEVATOR				
Present	2034	80%	392	74%
Not present	502	20%	141	26%
FURNISHED				
Yes	1649	65%	49	9%
No	71	3%	229	43%
Partially	814	32%	255	48%
N/C	2	0%	0	0%
TERRACE				
Present	925	36%	208	53%
Not present	1611	64%	253	47%
TYPE OF HOUSING				
House	0	0%	19	4%
Terraced house	31	1%	9	2%
Apartment	2505	99%	505	95%

Own elavoration

5. DATA

This section is dedicated to explain the adjustments made to the data in order to be prepared for the subsequent models:

[3] Houses include detached and semi-detached houses

The **dependent variable** (Rental price) has a natural logarithm form as well as the variable square meters, getting a **log-log** functional form between this two variables, the interpretation between them is given in % and represent the elasticity. This decision follows (*Yazdani, 2021*) work. It helps for an easier interpretation of the results as well as an attempt to try and reduce heteroscedasticity in the models. For the other variables a **log-linear** form is established. The dummy variables are interpreted such that they take the value 1 when present and 0 if not present.

In order to avoid exact collinearity, one of the two areas to be treated remains out of the equation. Additionally, and for the same reason, the variable "**Furnished**" has been modified into a dummy variable such that partially furnished dwellings have been considered as "Not furnished", taking a value 0.

A new variable named "**Luxury**" has been created by grouping the following characteristics: Garden, Terrace and Garage. This variable represents extra commodities that some dwellings include. It is specified as a dummy (taking value 1 when some of these characteristics are found).

There were some **missing/empty data** which could not be filled up, in order to deal with it (*Yazdani, 2021*) and (*J.Hill, 2013*) criteria has be established, "*numerical missing variables were filled with its mean and dummy variables with their mode*".

6. METHODOLOGY AND MODELS:

The methodologies used for model estimations are a multiple regression using OLS, Two Stage Least Squares (2SLS) and Generalized Methods of Moments (GMM) estimators with a cross-sectional data sample.

After performing the necessary tests, an endogeneity problem was detected, this is very common when using OLS in spatial models as it is often inconsistent, biased and no longer the best estimator with the presence of omitted variables. This is why, instrumental variables (IV) have been used by estimating with 2SLS and GMM. These two methods work with instrumental variables that help and give an effective explanation to the study variable as well as avoid endogeneity problems.

The variables "Rooms" and "Location" work as instruments for "Lsqm" and "Luxury", respectively. Both instruments have been checked to follow the two requisites to be a good IV, they are exogenous and partially correlated with the explanatory endogenous variable (*Wooldridge, 2008).* This IV have been selected because rooms and square meters explain the same thing measured in two different ways so when including both, the model is very likely to suffer from endogeneity. Same explanation is considered for the variable Luxury which is likely to have a high relation with Location.

a. MODELS:

There are three **models** to be tested:

First, a model in which City-Madrid and South-Madrid area are analyzed. In the second and third model, neighborhoods have been regrouped in order to test empirically what are the differences in the effect over the final rental price in the university areas and non-university areas.

$$(1)\ Lprice = \beta_0 + \beta_1 \cdot Lsqm + \beta_2 \cdot Bathrooms + \beta_3 \cdot SOUTH + \beta_4 \cdot AC + \beta_5 \cdot Pool + \beta_6 \cdot Elevator + \beta_7 \cdot Floor + \beta_8 \cdot Furnished + \beta_9 \cdot Terraced\ House + \beta_{10} \cdot Luxury + u$$

$$(2)\ Lprice = \beta_0 + \beta_1 \cdot Lsqm + \beta_2 \cdot Bathrooms + \beta_3 \cdot SOUTH\ UNIVERSITY\ AREA + \beta_4 \cdot AC + \beta_5 \cdot Pool + \beta_6 \cdot Elevator + \beta_7 \cdot Floor + \beta_8 \cdot Furnished + \beta_9 \cdot Terraced\ House + \beta_{10} \cdot Luxury + u$$

$$(3)\ Lprice = \beta_0 + \beta_1 \cdot Lsqm + \beta_2 \cdot Bathrooms + \beta_3 \cdot SOUTH\ NON\ UNIVERSITY\ AREA + \beta_4 \cdot AC + \beta_5 \cdot Pool + \beta_6 \cdot Elevator + \beta_7 \cdot Floor + \beta_8 \cdot Furnished + \beta_9 \cdot Terraced\ House + \beta_{10} \cdot Luxury + u$$

Models (2) and (3) follow the investigation line of (Jimoh et al., 2013). In both, a regrouping of the neighborhoods has been done. The selection criteria of university areas work as follows: In the South, the municipalities with public university (Mostoles, Alcorcon, Getafe, Fuenlabrada and Leganes) have been grouped. In the City, the neighborhoods with the highest registered number of potential university students (between 20 and 24 years old) have been selected, these are found to be living in the neighborhoods of Universidad, Embajadores, Rios Rosas and Gaztambide.

b. TESTS:

Once the models have been presented, several tests have been done in order to check the correct specification of the models as well as the robustness of the instruments used.

For the models using OLS, **residuals** were plotted, **correlations** were examined as well as the multicollinearity **test** (VIF) and the **homoskedasticity** white test. None of the models presented multicollinearity but all presented heterokedasticity, therefore following (*Wooldridge, 2008*) and (W*hite, 1980)* **robust standard errors** have been calculated in all regressions.

When using GMM and 2SLS for the estimations, the following tests have been calculated:

a. Endogeneity test: For this, Wu-Hausman, Durbin and GMM statistic tests were calculated. All of them reject the null: Variables are exogenous, therefore the use of 2SLS and GMM is justified as having endogenous regressor will cause OLS to fail.

b. Weakness instruments test, the first stage test was made in order to see if the instruments used to explain a high variation of the endogenous regressor.

Results show a rejection of H0: Instruments are weak; this leads us to conclude that there is no evidence of the weakness of the instruments used.

In order to see all tests and plots see APPENDIX B.

7. RESULTS

This section is dedicated to give the results of the estimated models:

a. SOUTH AND CITY MODEL

Table 4: Results of model (1), South vs City.

	OLS		2SLS		GMM	
Variable	Coeff	t	Coeff	z	Coeff	z
Lsqm	0.559***	30.8	0.452***	9.12	0.452***	8.1
Bathrooms	0.121***	9.27	0.145***	6.38	0.145***	5.75
SOUTH	-0.625***	-49.55	-0.699***	-30.54	-0.699***	-32.97
AC	0.147***	13.04	0.127***	8.46	0.127***	8.73
Pool	0.026	1.21	0.064**	2.42	0.064**	2.43
Elevator	0.062***	4.4	0.045**	2.53	0.045**	2.51
Floor	0.009**	3.08	0.006**	2.29	0.007**	1.98
Furnished	-0.115***	-5.04	-0.076**	-2.13	-0.076**	-2.56
Terraced House	-0.036	-1.02	-0.016	-0.27	-0.016	-0.33
Luxury	-0.004	-0.41	0.438***	5.41	0.438***	5.28
Cons	4.571***	72.12	4.87***	29.83	4.87***	26.38
R-Squared	69.72%		55.43%		55.43%	
Statistic	F-stat (10,3058)	816.04	Wald chi2 (10)	4193.64	Wald chi2 (10)	4445.54
Obs	3,069		3,069		3,069	

Note: Coefficients of the variables for the different regression with Lprice as dependent variable. All standard deviations are calculated as ***robust*** *because of previously checked heterocedasticity*** p<0.01. ** p<0.05. * p<0.1.*

In general, all variables follow the current literature expectations (Table 1). In model (1) all variables are found to be statistically significant except for Terraced Houses, a possible explanation for this fact is that there are very few Terraced Houses in the sample. The goodness of fit or R^2 gets a higher value than 55% meaning more than 55% of the rental price is explained by the independent variables included in the model. The relevant variable in this model, “SOUTH”, presents a negative coefficient. It is estimated that the rent of a house with the same characteristics located in the **South** is nearly **70% cheaper** than in the City. Having an extra bathroom in the dwelling produces a substantial increase in the final rental price (14,5%).

Significance and coefficients between 2SLS and GMM do not vary much, this makes us think that endogeneity problems have been fixed. The variable “Luxury” changes from been not significant when using OLS to highly significant when using 2SLS and GMM, this could infer that the instrument added helps the explanation of this variable. This variable is important to comment due to the time frame in which the data has being collected (post-pandemic era), it shows it is estimated an increase in the rental price of nearly 44% when presenting Terrace, Garden or Garage.

Big differences between variables can be seen, it is estimated that a 1% increase in the square meters leads to an increase of approximately 45% of the rental price while “Floor” and “Furnished” have a very low impact on reducing the final rental price in 0.6% and 7%, respectively.

b. UNIVERSITY AND NON-UNIVERSITY AREA MODELS

Table 5: Results of model (2), University area.

	OLS		2SLS		GMM	
Variable	Coeff	t	Coeff	z	Coeff	z
Lsqm	0.531***	20.15	0.444***	5.17	0.444***	5.03
Bathrooms	0.124***	6.69	0.131***	3.28	0.131***	3.33
SOUTH UNIVERSITY AREA	-0.573***	-34.1	-0.754***	-14.96	-0.754***	-15.66
AC	0.112***	6.78	0.102***	4.9	0.102***	4.95
Pool	0.054**	2.09	0.158***	3.53	0.158***	3.89
Elevator	0.0037*	1.84	0.027	1.04	0.027	1
Floor	0.016***	3.25	0.014***	2.79	0.014**	2.24
Furnished	-0.011***	-3.23	-0.047	-0.78	-0.047	-0.98
Terraced House	-0.116***	-3.31	-0.010	-1.1	-0.010	-1.6
Luxury	-0.011	-0.65	0.470***	3.68	0.470***	3.71
Cons	4.655***	52.94	4.873***	18.02	4.873***	17.35
R-Squared	66.63%		47.78%		47.78%	
Statistic	F-stat (10,1242)	386.24	Wald chi2 (10)	1463.01	Wald chi2 (10)	2097.2
Obs	1,253		1,253		1,253	

*Note: Coefficients of the variables for the different regression with Lprice as dependent variable. All standard deviations are calculated as **robust** because of previously checked heterocedasticity*** p<0.01. ** p<0.05. * p<0.1.*

Table 6: Results of model (3), Non-University area.

	OLS		2SLS		GMM	
Variable	Coeff	t	Coeff	z	Coeff	z
Lsqm	0.573***	23.68	0.444***	8.5	0.444***	8.67
Bathrooms	0.115***	6.93	0.156***	6.37	0.156***	6.61
NON UNIVERSITY SOUTH AREA	-0.767***	-31.63	-0.841***	-23.65	-0.841***	-21.23
AC	0.172***	10.7	0.147***	7.19	0.147***	6.87
Pool	0.082**	2.01	0.115**	2.44	0.115***	2.82
Elevator	0.093***	4.37	0.081***	3.27	0.081***	3.24
Floor	0.005	1.57	0.004	1.01	0.004	0.98
Furnished	-0.121***	-3.97	-0.081**	-2.16	-0.081*	-1.79

Terraced House	0.006	0.14	0.049	0.88	0.049	0.66
Luxury	-0.002	-0.14	0.394***	3.95	0.394***	4
Cons	4.516***	50.82	4.91***	27.9	4.91***	28.55
R-Squared	69.09%		58.48%		58.48%	
Statistic	F-stat (10,1698)	481.99	Wald chi2 (10)	2825.2	Wald chi2 (10)	2447.34
Obs	1,709		1,709		1,709	

Note: Coefficients of the variables for the different regression with Lprice as dependent variable. All standard deviations are calculated as ***robust*** *because of previously checked heterocedasticity*** p<0.01. ** p<0.05. * p<0.1.*

Regarding the university areas, independently of whether it is in the university or non-university neighborhood the southern zone has a lower price. This result reinforces the conclusions of the model (1). However, it is relevant to comment that the rental price differential between the South and the City is greater in non-university residential areas than in university ones.

Comparing models (2) and (3) relevant results can be highlighted. Having Pool is more valued in the university areas rather than the non-university, increasing the price from 15% in the university areas to 11% in the non-university. It is estimated that dwellings in university areas increase 10% their rental price when having air conditioners whereas in the non-university areas this increase is of 15%. A similar situation occurs when having Elevator, which increases the price from 2% to 8%, respectively. Talking about the variable “Luxury” (having a terrace, garage or garden), it is estimated to increase the final rental price in a higher percentage in university areas than in non-university areas (57% and 38%, respectively).

Regarding the R^2 the independent variables included in the models explain 47% of the university area rental prices and 58% of the non-university area prices. This difference, although small, could be explained as there might be some characteristics not included in the model which are valued in university areas such as distance to the public transport which are not taken into account in this analysis.

8. CONCLUSIONS

Given the post-pandemic era of the data collected, our living conditions have changed and the importance of choosing the correct home to reside has increased. This fact has affected the value and usefulness that we have given to homes and their characteristics. Over the years, it could be observed how, given the needs of the population, the percentage weight on the price of the characteristics of a rental home varies. Thinking about the results obtained during the analysis it is important to highlight the variable Luxury which increases the rental price by approximately 40%.

Housing is a relevant asset for the well-being of society, that is why carrying out a correct analysis of it is very important. Since the introduction of the Hedonic price model in the Housing market, it has been a very used method to explain prices all around the world as it "*measures the implicit price of each attribute*" *(O´Sullivan, 2011).*

One of the complexities of this analysis is the difficult availability of the data. In this project, the collection of data was made from the main web portal of buy, sale and rent advertisements in the country, "*El Ideaslista*" during the time period of early 2022. The database is formed by 3.069 rental houses in the areas of Madrid City and South Madrid. Several models have been regressed using OLS, 2SLS and GMM to solve endogeneity problems.

The results obtained are in line with the general reviewed literature. The main points to highlight from this econometric study are the following: it is shown that the same house located in the **suburbs is a 65% cheaper**. Between the characteristics that cause the greatest percentage increase in the rental price are having one more bathroom and Luxury which imply having terrace, garden or garage when renting a house.

Regarding the areas nearby the main universities of the City and South suburbs (**university models)** that have been estimated, the most remarkable result is that a rental house with Pool, Bathroom or Luxury characteristics increase more the rental price in University areas than in non-university areas. Another remarkable conclusion from this models is the gap differential. The difference in rental prices is smaller between the university areas than between the non-university areas, this analysis strengthens the previous result in which the South is cheaper in all areas compared to the City.

A way of **improving this analysis** could be increasing the collection period. Ours conclusions have been reached for the limited period of data collection (February to April), so extending the data collection period could lead to higher robust results. The variable "Year" has not been added to the models as currently, it was not possible to find data on it, another way of improving it could be by finding this variable and watching if the number of years a house has affects the total rental price. Distance to the public transport or distance to the university could also be relevant variables to include in further analysis. As seen in previous literature using a time series with periodical cross-sectional data could also be interesting as dynamics can be analyzed and are important in order to see how prices in different areas react to business cycles.

9. BIBLIOGRAPHY

1. *Articulo 47, Capítulo III: De los principios rectores de la política social y económica.Título I de los Derechos Fundamentales y Deberes* 1978, , Artículo 47,. Capítulo tercero.

2. «BOE» núm. 117 & Ministerio de Sanidad y Consumo 1989, *Real decreto 515/1989, sobre protección de los consumidores en cuanto a la información de suministrar en la compra-venta y arrendamiento de viviendas*.

3. Ann D. Witte, Howard J. Sumka and Homer Erekson 1979, "An Estimate of a Structural Hedonic Price Model of the Housing Market: An Application of Rosen's Theory of Implicit Markets ", vol. 47, no. 5, pp. 1151-1173.

4. Anne Sanftenberg 2018, "Specification of a Hedonic Model for Rental Apartment Buildings in Berlin", *Journal of applied Business and Economics,* vol. 20, no. 9.

5. Arthur O´Sullivan 2011, *URBAN ECONOMICS* , 8th edition edn, McGraw-Hill/Irwin.

6. Ayuntamiento de Madrid 2022, , *Padrón Municipal. Alta y cambio de domicilio en Padrón*. Available: https://sede.madrid.es/portal/site/tramites/menuitem.62876cb64654a55e2dbd7003a8a409a0/?vgnextoid=3e3debb41f6e2410VgnVCM2000000c205a0aRCRD&vgnextchannel=775ba38813180210VgnVCM100000c90da8c0RCRD&vgnextfmt=default.

7. Comunidad de Madrid 2009, *RELACIÓN DE MUNICIPIOS POR ÁREAS TERRITORIALES* , Portal Web del Ayuntamiento de Madrid.

8. Desormeaux Daniela & Piguillem Facundo 2003, *Precios Hedónicos e Índices de Precios de Viviendas*.

9. Dimitra Kavarnou & Anupam Nanda 2015, "HOUSE PRICE DYNAMICS IN PANAMA CITY", *Journal of Real Estate Literature,* vol. 23, no. 2, pp. 315-334.

10. Dimitra Kavarnou & Anupam Nanda 2014, "Which Attributes Influence the Housing Markets across the Greek Islands?", *Journal of Real Estate Literature,* vol. 22, no. 2, pp. 233-260.

11. Dirección General de Estadística 2017, Noviembre-last update, *Mapa Distritos de Madrid - Distrito 01- Centro*. Available: https://www.madrid.es/portales/munimadrid/es/Inicio/El-Ayuntamiento/Estadistica/Areas-de-informacion-estadistica/Territorio-climatologia-y-medio-ambiente/Territorio/Mapas-de-Distritos-y-Barrios/?vgnextfmt=default&vgnextoid=240d64c49579f410VgnVCM1000000b205a0aRCRD&vgnextchannel=e59b40ebd232a210VgnVCM1000000b205a0aRCRD.

12. Embaye, W.T., Zereyesus, Y.A. & Chen, B. 2021, "Predicting the rental value of houses in household surveys in Tanzania, Uganda and Malawi: Evaluations of hedonic pricing and machine learning approaches", *PloS one,* vol. 16, no. 2, pp. e0244953.

13. Ha Minh NGUYEN , Hung Quoc PHAN , Tri Van TRAN, Thang Kiem Viet TRAN 2020, *The Hedonic Method in Evaluating Apartment Price: A Case of Ho Chi Minh City, Vietnam*.

14. Hanink, D.M., Cromley, R.G. & Ebenstein, A.Y. 2010, "Spatial Variation in the Determinants of House Prices and Apartment Rents in China", *The journal of real estate finance and economics,* vol. 45, no. 2, pp. 347-363.

15. Jeffrey M. Wooldridge 2001, "Applications of Generalized Method of Moments Estimation", *Journal of Economic Perspectives,* vol. 15, no. 4, pp. 87-100.

16. Jeffrey M.Wooldridge 2008, *Introduction to Econometrics,* 4th edn, Cengage Learning.

17. Jimoh, S., Abdulrazaq, B., Umar, I. & Sulaiman, L.A. 2013, *AN ECONOMIC ANALYSIS OF DETERMINANTS OF HOUSE RENTS IN THE UNIVERSITY ENVIRONMENT*.

18. Kelvin J. Lancaster 1966, "A New Approach to Consumer Theory", *Journal of Political Economy*, vol. 74, no. 2, pp. 132- 157.

19. Manuel A. Zambrano-Monserrate 2015, *Formación de los precios de alquiler de viviendas en Machala (Ecuador): análisis mediante el método de precios hedónicos*, 2015 Asociación Cuadernos de Economía. Publicado por Elsevier España, S.L.U.

20. Maslow, A.H. 1943, "A theory of human motivation." in *50(4), 370–396*Psychological Review.

21. My-Linh 2020, "The Hedonic Pricing Model Applied to the Housing Market", *International Journal of Economics and Business Administration,* vol. VIII, no. Issue 3, pp. 416-428.

22. Raya Vílchez, J.M. 2005, *Ensayos sobre el mercado de la vivienda en España: Precios hedónicos y regímenes de tenencia*, Universitat de Barcelona.

23. Robert J.Hill 2013, "Hedonic price indexes for residential housing: a survey, evaluation and taxonomy", *Journal of Economic Surveys, University of Graz,* vol. 27, no. 36-2011/1, pp. 879-914.

24. SELİM Sibel 2008, *DETERMINANTS OF HOUSE PRICES IN TURKEY: A HEDONIC REGRESSION MODEL* .

25. Sherwin Rosen 1974, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition ", *The University of Chicago Press,* vol. 82, no. 1.

26. Similar web 2022, March-last update, *Top Websites Ranking*. Available: https://www.similarweb.com/top-websites/spain/category/business-and-consumer-services/.

27. Sirmans Stacy, G., Macpherson David, A. & Zietz Emily, N. 2005, *The composition of Hedonic Pricing Models*.

28. Solovev Kirill & PröllochsNicolas 2021, *Integrating Floor Plans into Hedonic Models for Rent Price Appraisal*.

29. Tung-Leong, C. & K W CHAU* 2003, "A Critical Review of Literature on the Hedonic Price Model", *International Journal for Housing and Its Applications,* vol. 27(2), pp. 145-165.

30. Waddell, P., Berry, B.J.L. & Hoch, I. 1993, "Residential property values in a multinodal urban area: New evidence on the implicit price of location", *The journal of real estate finance and economics,* vol. 7, no. 2, pp. 117-141.

31. Wilson, B. & Frew, J. 2007, *Apartment Rents and Locations in Portland, Oregon*.

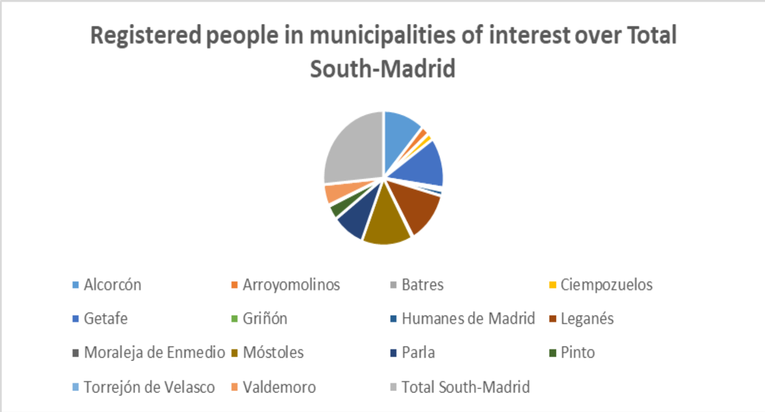
32. Xiaochen H, Luechai Chulasai, Sasipen Phuangsaichai 2011, "Hedonic Pricing Model for Housing Market in City of Kunming, the People's Republic of China", *CMU. JOURNAL OF ECONOMICS,* pp. 145-165.

33. Yazdani, M. 2021, "House Price Determinants and Market Segmentation in Boulder, Colorado: A Hedonic Price Approach".

34. Zakarina Firano, Fatine A. Filiali 2021, *Towards the hedonic modelling and determinants of real estates price in Moroco*, Elsevier.

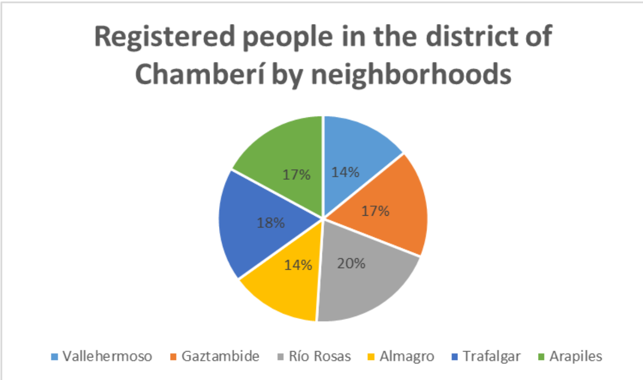
APPENDIX A

Table 7: Registered people in municipalities of interest over Total South-Madrid



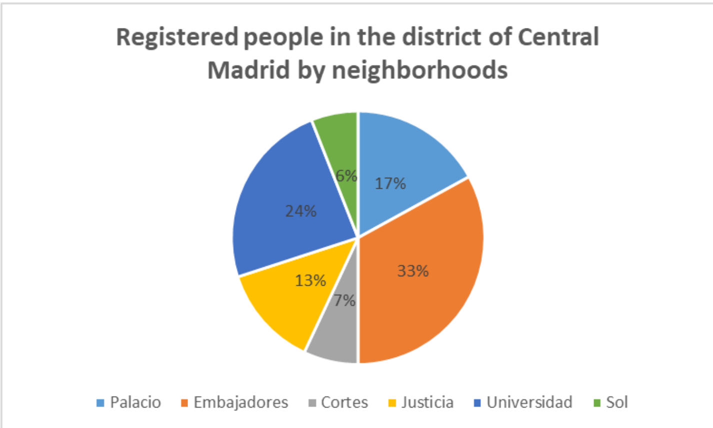
Own elaboration based on (Ayuntamiento de Madrid, 2022)

Table 8: Registered people in the district of Chamberí by neighborhoods



Own elaboration based on (Ayuntamiento de Madrid, 2022)

Table 9: Registered people in the district of Central Madrid by neighborhoods



Own elaboration based on (Ayuntamiento de Madrid, 2022)

APPENDIX B

Table 10: VIF and Residuals Plot for Model (1)

VIF		
Variable	*VIF*	*1/VIF*
Lsqm	*2.34*	*0.4274*
Bathrooms	*2.18*	*0.4585*
SOUTH	*1.24*	*0.8071*
Pool	*1.20*	*0.8343*
AC	*1.10*	*0.9060*
Elevator	*1.10*	*0.9099*
Luxury	*1.09*	*0.9184*
Floor	*1.05*	*0.9490*
Furnished	*1.03*	*0.9696*
Terraced House	*1.01*	*0.9950*
Mean VIF	*1.33*	*-*

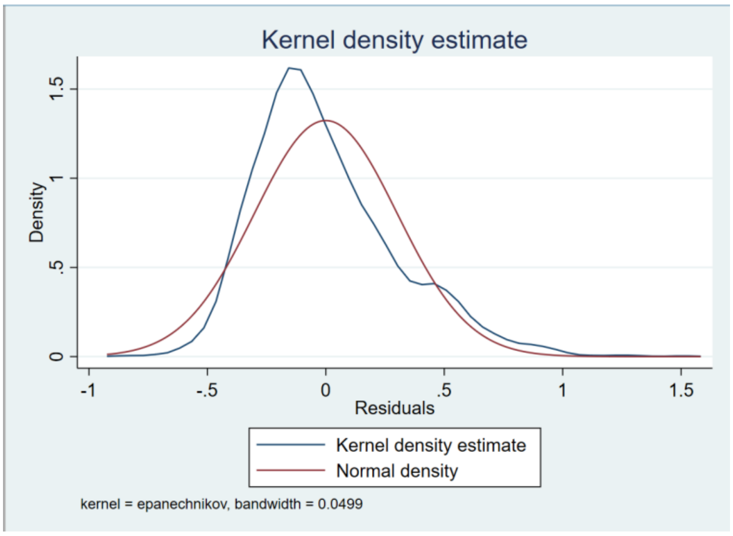


Table 11: Test of endogeneity for Model (1)

Test of endogeneity

GMM	*GMM C statistic*	*chi2(2) = 47.5724*	*(p = 0.0000)*
2SLS	*Durbin (score)*	*chi2(2)= 53.4527*	*(p = 0.0000)*
	Wu-Hausman	*F(2,3056)= 27.0849*	*(p = 0.0000)*

Note: H0, Variables are exogenous

Table 12: Weakness of instruments test for Model (1)

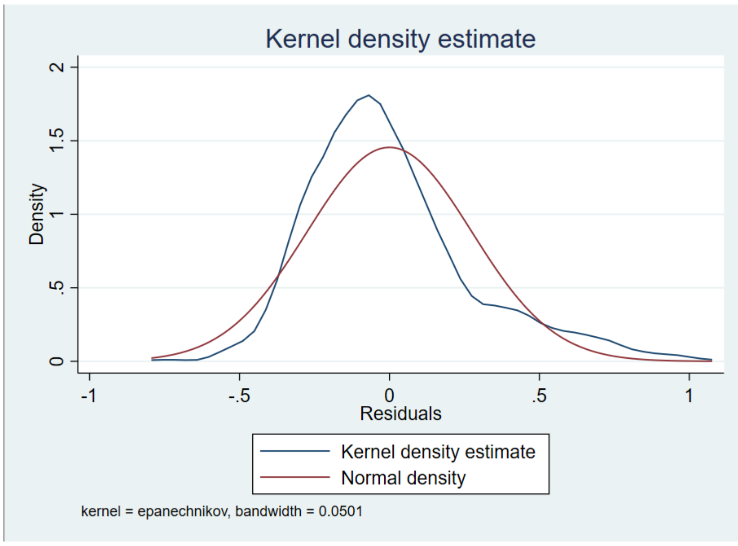
Firststage test:

Minimum eigenvalue statistic = 47.7742				
	10%	*15%*	*20%*	*25%*
2SLS Size of nominal 5% Wald test	*7.03*	*4.58*	*3.95*	*3.63*

Note: H0, Instruments are exogenous

Table 13: VIF and residuals plot for Model (2)

VIF		
Variable	*VIF*	*1/VIF*
Lsqm	*2.50*	*0.3995*
Bathrooms	*2.28*	*0.4381*
SOUTH UNI AREA	*1.32*	*0.7570*
Pool	*1.34*	*0.7448*
AC	*1.08*	*0.9241*
Elevator	*1.13*	*0.8848*
Luxury	*1.22*	*0.8230*
Floor	*1.07*	*0.9389*



Furnished	*1.02*	*0.9842*
Terraced House	*1.02*	*0.9851*
Mean VIF	*1.40*	*-*

Table 14: Test of endogeneity for Model (2)

Test of endogeneity

GMM	*GMM C statistic*	*chi2(2) = 37.1116*	*(p = 0.0000)*
2SLS	*Durbin (score)*	*chi2(2)= 41.9643*	*(p = 0.0000)*
	Wu-Hausman	*F(2,3056)= 21.484*	*(p = 0.0000)*

Table 15: Weakness of instruments test for Model (2)

Firststage test:

Minimum eigenvalue statistic = 18.1195				
	10%	*15%*	*20%*	*25%*
2SLS Size of nominal 5% Wald test	*7.03*	*4.58*	*3.95*	*3.63*

Table 16: VIF and residuals plot for Model (2)

VIF		
Variable	*VIF*	*1/VIF*
Lsqm	*2.30*	*0.4349*
Bathrooms	*2.17*	*0.4599*
NON SOUTH UNI AREA	*1.36*	*0.7352*
Pool	*1.29*	*0.7775*
AC	*1.10*	*0.9114*
Elevator	*1.06*	*0.9449*
Luxury	*1.07*	*0.9311*
Floor	*1.05*	*0.9563*
Furnished	*1.07*	*0.9364*
Terraced House	*1.01*	*0.9932*
Mean VIF	*1.35*	*-*

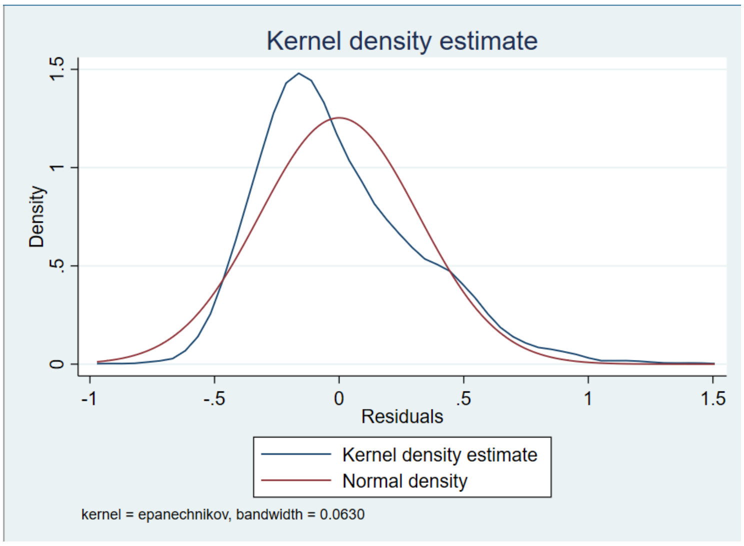


Table 17: Test of endogeneity for Model (3)

Test of endogeneity

GMM	*GMM C statistic*	*chi2(2) = 21.1604*	*(p = 0.0000)*
2SLS	*Durbin (score)*	*chi2(2)= 22.7238*	*(p = 0.0000)*
	Wu-Hausman	*F(2,3056)= 11.4274*	*(p = 0.0000)*

Table 18: Weakness of instruments test for Model (3)

Firststage test:

Minimum eigenvalue statistic = 32.7234				
	10%	*15%*	*20%*	*25%*
2SLS Size of nominal 5% Wald test	*7.03*	*4.58*	*3.95*	*3.63*