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

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 variables affecting the rental house market in Madrid City and South Madrid has been carried out, using a hedonic price model. Landlords and tenants can use this approach in order to establish the rental price of their livings based on the variables that most affect the final rental price.

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 counts with a lower rental price in the City whereas the university areas located in the South suffer a higher rental price.

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 possible.

Housing is one of the biggest investments during the life of a citizen and, as indicated in the real decree 515/1989, "*Housing is currently one of these products of ordinary and widespread use. Its use through purchases or lease,* ***constitutes an activity not only daily, but of great importance in the life of a consumer***" *(«BOE» núm. 117, Ministerio de Sanidad y Consumo, 1989).*

Houses are assets which are not only needed for "***Safety***" (*Maslow, 1943*) but also used as a **speculative and investment goods,** in particular when talking about renting. In this case, landlords, usually already owners of another house, make the investment decision on buying a second one in order to receive some revenues of it. In order to evaluate and estimate the return of this investment, it is important to look at the renting price, which is determined by some specific characteristics or variables of the property.

Some of these variables are intrinsic to the home itself and others depend on the environment and location. One of the main variables which have an impact on the final renting price is the area where the property is located, especially in big cities like Madrid, London or Berlin where prices vary a lot among different areas. This is the reason of why previous researches have focus their analysis in areas of Portland (*Wilson, Frew, 2007*), 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 analyze also other different variables through the *Hedonic Price Models.*

Hedonic Price Model estimates the value of the variable to be analyzed (in this project, the rental price), through the sum of explanatory variables. Historically, hedonic price models were introduced into the housing market firstly by *(Lancaster, 1966)* and more deeply by *(Rosen, 1974).* This model allows to estimate the impact of the characteristics of the house in the renting price. Besides the location, other relevant variables are square meters, rooms, bathrooms and other dummy variables such as terrace or air conditioner. In other words, this model helps us explain which variables are more valued to define the rental price in the different housing markets.

Understanding the concrete renting market of each area is important to **policy makers**, **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 air conditioner is installed in the house? From the demander point of view, it provides an empirical evidence about the

expected rent 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 policy makers in order to better understand and regulate the market in the studied areas.

This project analyzes the **rental price** in the areas of **City** and **South Madrid,** which according to literature, have not been analyzed up to now. For the evaluation and models, actual and current market data has been considered, collecting more than 3000 entries from the main housing supply platform in Spain "*El Idealista*[1]" which form this project database.

The **main target** is to find which commodities affect the most the rental prices, which variables are not significant and what is the difference between living in City or South Madrid.

The **methodology** used in this economic analysis with econometric support is a hedonic price model, widely used for this economic works. In order to achieve the models, multiple linear regressions are calculated using the Ordinary least squares (**OLS**) estimator, following the steps from *(Wooldridge, 2008).*

As a final summary, Madrid has a very heterogeneous housing market, with a wide range of different supplies, commodities, prices and locations. The aim of this study is to contribute to the current literature by analyzing the rental prices of Madrid City and South Madrid through a hedonic price model, considering different explanatory variables.

1. OBJECTIVES AND STRUCTURE:

The main objective of this final bachelor project is to identify the most significant variables which fix the rental price in the areas of Central Madrid and South Madrid applying and 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 renting 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 acquire data.

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

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 literature review, section 4 explains data as well as the areas of interest, section 5 analyzes the different Hedonic price models to be analyzed, section 6 shows the models results as well as the intuitions generated from them. Finally, chapter 7 summarizes the most important final conclusions.

2. LITERATURE REVIEW

There are many previous case studies which use the hedonic pricing model in their papers. This section gives and overview of the previous literature and research on this topic, which has helped the current work on the selection of the variables to include in the model, as well as in the general understanding of the topic.

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

Later in time, *(Tung-Leong, K W CHAU*, 2003)* made a literature review of previous studies in which several conclusions can be made. First of all, this author emphasized the difficulty on making this type of analysis as they are "*Constrained by the availability of data*" and described it as the "*major scientific method by which we can observe the effects of one or more housing attributes on housing prices*". Not only this but this paper summarizes and characterizes the housing attributes in 3 sections: Locational, Structural and Neighborhood characteristics and collected the general sign (positive or negative) the previous literature found about the variables.

Table 1 is important in order to give a previous intuition in what could the sign of the variables of the current work look like, according to the previous literature findings.

"*Homes are heterogeneous goods, the environment and the consumer vary in each area*" (*Stacy, Macpherson & Zietz, 2005)* but there are several common and important characteristics that explain the price of a house which should be analyzed in all models. Commonly in all case studies square meters, number of rooms and number of bathrooms are included in the model regardless the location. In contrast with this, there are several studies in Latin American which add variables that may not be considered for other countries given the level of development and accessibility to them. Some examples of this cases are, water supply (permanent or irregular) (*Zambrano-Monserrate, 2015*) or access to hot water or cable TV (*Sibel, 2008*).

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*

(Stacy, Macpherson & Zietz, 2005) paper also 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.

The authors (*Dimitra Kavarnou and Anupam Nanda*) joint several times to analyze two different areas. The first one *(Kavarnou, Nanda, 2014)* did a hedonic price model for the Greek Islands. It is a very unique study as this area was mostly not covered before, they used cross-sectional data for their analysis meaning they were able to capture the dynamics of the housing market. The results follow the common literature. For example, the living space and land area showed a positive impact in 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 (*Kavarnou, Nanda, 2015*) Panama´s housing market was investigated. They divided the country into 11 zones, quarterly data since 2005 to 2014 was involved in the analysis. A positive and highly statistically significant coefficient for the variables "Bathrooms", "Ln(sqm[2])", "Service_room" were shown during the estimation whereas "No covered parking" showed a negative coefficient.

[2] Sqm refers to square meters (m^2)

Regarding the appraisal of goods, despite the fact that this work focuses on renting them, it 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 think what other variables could be relevant, new and valued when renting or buying a house which are not so intuitive to explain housing prices.

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.

3. AREAS AND DESCRIPTIVE STATISTICS:

4.1 AREAS:

Madrid´s community is made up by 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.

Talking about South-Madrid, it is composed by 39 municipalities but the only ones with supply of rent 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.

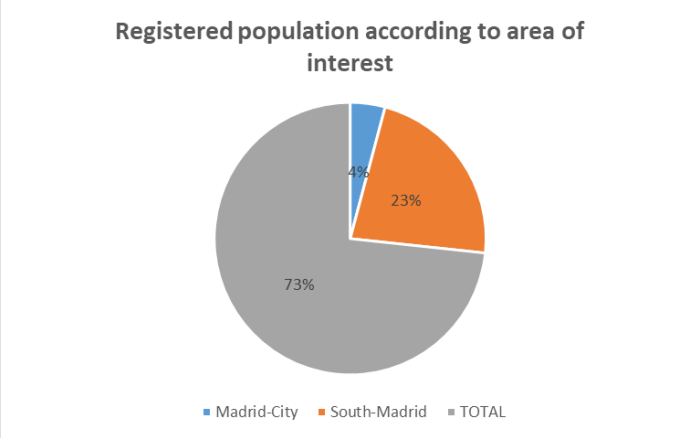
Prior to the model and analysis of the database, an analysis of the size of the zones in terms of their population will be done taking as reference the number of registered residents in each area. The selection of registered residents is taken as reference given that the Madrid city council states "*Anyone who lives in Spain is **obliged to register in the municipal register where they habitually reside**. Whoever lives in more than one municipality will register in the one they reside for the longest time per year"*. Therefore, it is understood, that the number of registered people represent 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 people. Madrid-City constitutes a 4% of the total with 281.358 people whereas South-Madrid represents the 23% of the total population with 1.527.838

people. *(Ayuntamiento de Madrid, 2022).* Therefore, there are more people living in the South rather than in the City area.

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 neighborhood with more registered people in Central-Madrid is Embajadores (33% of the total) whereas for Chamberí this neighborhood is Río Rosas (20% of the total) while, 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)

4.2 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 to extract the data and not another real estate web portal? The main answer is: because it is the leading web portal for purchase and rental offers in Spain. According to *(Similar web, 2022)* is has the number one position in 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 have been collected in Madrid-City (1.586 data for the Cental district and 950 for Chamberí district) 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
Storager	Storage Room	Cualitative variable, dummy
Pool	Pool	Cualitative variable, dummy
Elevator	Elevator	Cualitative variable, dummy
Furnished[3]	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

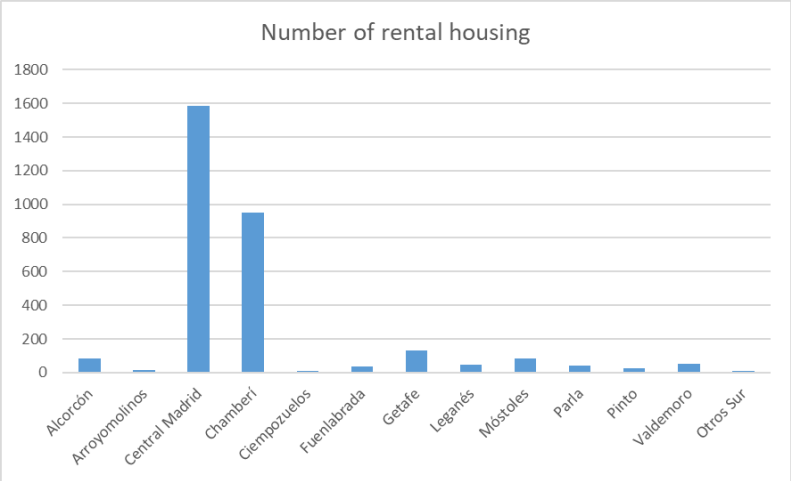
It is important to note that in order to elaborate the hedonic price model the variable "**District**" has been selected instead of "Distance to the center". This can also be seen in previous research (*Sanftenberg, 2018*), in the city of Berlin, because of the characteristics of the area, big polycentric cities like Madrid or Berlin do not count with a main center of reference therefore selecting the districts is a better way of classifying the data.

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

Figure 2 shows that the area with the most offer of rental houses is the City made up by the district of Central Madrid with nearly 1600 offers and followed by Chamberi. We can conclude that the City offers much more renting houses compared to the South.

[3] 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 2: Number of rental housing

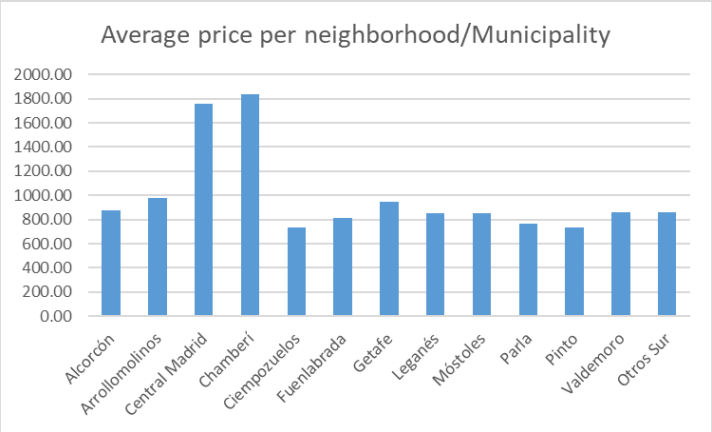


Own Elaboration

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.

Figure 3: Average price per neighborhood/municipality

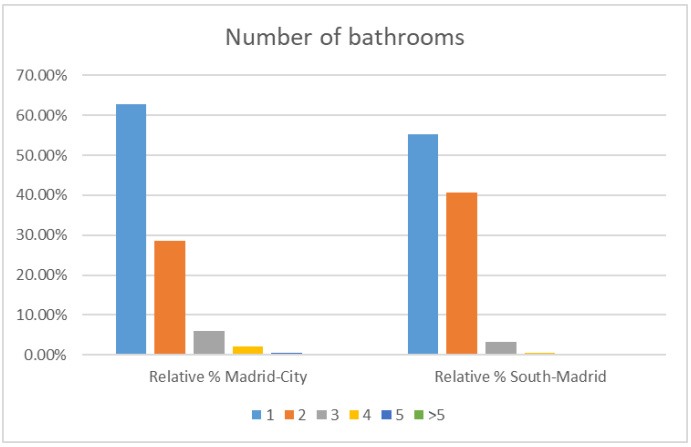


Own Elaboration

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

Regarding other characteristics such as the “Number of bathrooms” and the “Number of Rooms” the following is observed.

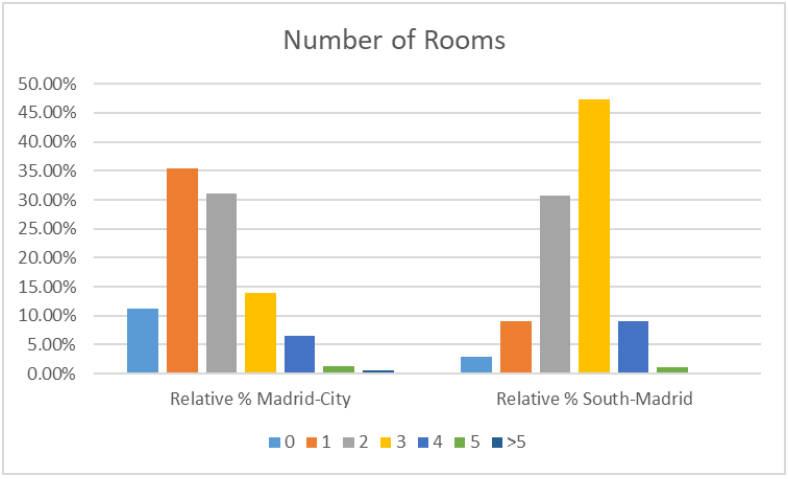
Figure 4: Number of bathrooms



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 this 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

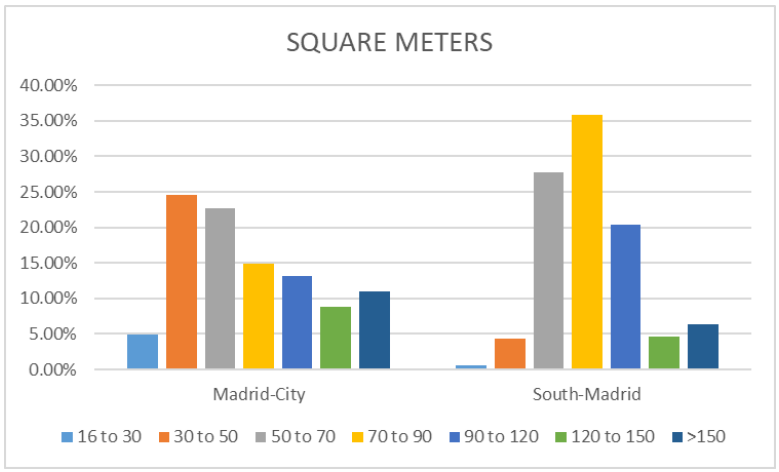


Own Elaboration

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

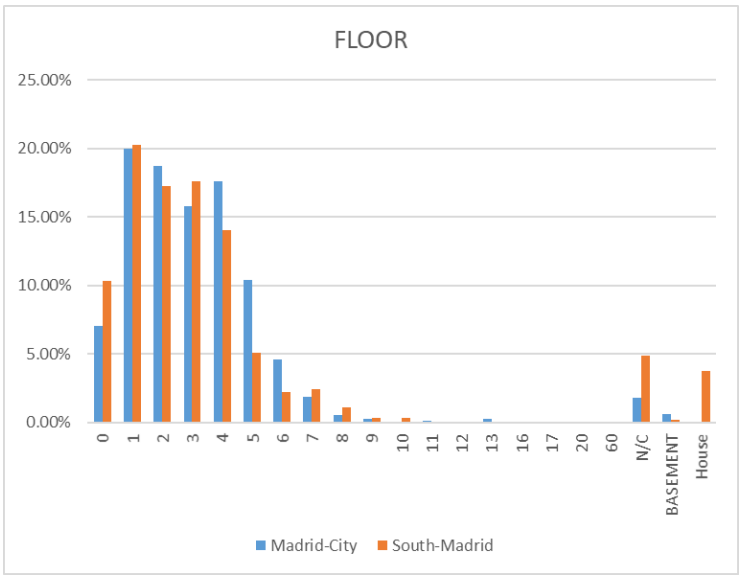
Supplied rental houses in the City commonly have between 30 and 50 square meters while the ones in the South tend to have between 70 to 90 square meters, as it is illustrated in Figure 6. Looking at some of the relations we can highlight the following: there are very few houses in the South with 16 to 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 us to conclude that houses for rent in the **South** tend to have **relatively more square meters than** the ones in the **City**. In conclusion, the **mode** in the **City** is to have between **30 to 50 sqm** whereas in the **South** the mode is to have between **70 to 90 sqm**.

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 floor**. Flats are only avaible 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:

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%
STORAGE ROOM				
Present	215	8%	154	29%
Not present	2321	92%	379	71%
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

It is observed that, in the City, 69% of the dwellings have **air condinitioner** whereas in the South less than 50% have it. Therefore it is very common to have this feature in both but it is less common in the South. On the other hand, 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.

Talking about **storage rooms** it is not common to have them when renting a house in none of the areas. 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** [4]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**.

4. HEDONIC PRICE MODELS AND METHODOLOGY:

Hedonic price models in the housing market "*measures the implicit price of each attribute*" *(O´Sullivan, 2011).*

The **endogenous 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, therefore 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 value 1 when present and 0 when 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 value 0.

[4] Houses include detached and semi-detached houses

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, "*to avoid losing the samples from analysis numerical missing variables were filled with its mean and dummy variables with their mode*".

Before running the regression, **residuals** where plotted, **correlations** were examined as well as the **multicolliearity test** (VIF) and the **homokedasticity** 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. (see Own elaboration based on (Ayuntamiento de Madrid, 2022)

APPENDIX B for all tests and plots)

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

First, a **general model** in which **City-Madrid and South-Madrid** area are analyzed. In the second model, neighborhoods have been regrouped in order to test empirically if the **university areas** have an impact and help to explain the rental price.

5.1 MODEL 1: CITY AND SOUTH MADRID

The first model its's made up of two regressions, all corrected for heterokedasticity with robust standard errors.

$$(1)\ Lprice = \beta_0 + \beta_1 \cdot Lsqm + \beta_2 \cdot Rooms + \beta_3 \cdot Bathrooms + \beta_4 \cdot SOUTH + u$$

$$(2)\ Lprice = \beta_0 + \beta_1 \cdot Lsqm + \beta_2 \cdot Rooms + \beta_3 \cdot Bathrooms + \beta_4 \cdot SOUTH + \beta_5 \cdot Terrace + \beta_6 \cdot AC + \beta_7 \cdot Garage + \beta_8 \cdot Storage + \beta_9 \cdot Pool + \beta_{10} \cdot Location + \beta_{11} \cdot Garden + \beta_{12} \cdot Elevator + \beta_{13} \cdot Floor + \beta_{14} \cdot Furnished + \beta_{15} \cdot Terraced\ house + \beta_{16} \cdot House + u$$

Regression (1) gives an output of the main/control variables included in most of the models according to the literature and regression (2) adds all other explanatory variables forming the database.

5.2 MODEL 2: UNIVERSITY MODEL

This model has been created in line with *(Jimoh et al., 2013).* We can distinguish this model into two, **Madrid-City university model** and **South-Madrid university model**.

In both, a regrouping of the neighborhoods has been done. In **Madrid-City** the University variable selection has been made by adding into this dummy variable the **neighborhoods with the highest registered number of potential university students** (between 20 and 24 years old). These are found to be living in the neighborhoods of Universidad, Embajadores, Ríos Rosas and Gaztambide so these 4 neighborhoods have been grouped into "City University area" variable. *(Ayuntamiento de Madrid, 2022)*

For the **South-University model** the **municipalities with public university** have been grouped into "South University area" variable. These municipalities are Mostoles, Alcorcon, Getafe, Fuenlabrada and Leganes.

All models are corrected for heterokedasticity with robust standard errors.

5.2.1 CITY-MADRID UNIVERSITY AREA:

$$(3)\ Lprice = \beta_0 + \beta_1 \cdot Lsqm + \beta_2 \cdot Rooms + \beta_3 \cdot Bathrooms + \beta_4 \cdot CITY\ UNIVERSITY AREA + u$$

$$(4)\ Lprice = \beta_0 + \beta_1 \cdot Lsqm + \beta_2 \cdot Rooms + \beta_3 \cdot Bathrooms + \beta_4 \cdot CITY\ UNIVERSITY AREA + \beta_5 \cdot Terrace + \beta_6 \cdot AC + \beta_7 \cdot Garage + \beta_8 \cdot Storage + \beta_9 \cdot Pool + \beta_{10} \cdot Location + \beta_{11} \cdot Garden + \beta_{12} \cdot Elevator + \beta_{13} \cdot Floor + \beta_{14} \cdot Furnished + \beta_{15} \cdot Terraced\ house + u$$

5.2.1 SOUTH-MADRID UNIVERSITY AREA:

$$(5)\ Lprice = \beta_0 + \beta_1 \cdot Lsqm + \beta_2 \cdot Rooms + \beta_3 \cdot Bathrooms + \beta_4 \cdot SOUTHUNIVERSITY AREA + u$$

$$(6)\ Lprice = \beta_0 + \beta_1 \cdot Lsqm + \beta_2 \cdot Rooms + \beta_3 \cdot Bathrooms + \beta_4 \cdot SOUTHUNIVERSITY AREA + \beta_5 \cdot Terrace + \beta_6 \cdot AC + \beta_7 \cdot Garage + \beta_8 \cdot Storage + \beta_9 \cdot Pool + \beta_{10} \cdot Location + \beta_{11} \cdot Garden + \beta_{12} \cdot Elevator + \beta_{13} \cdot Floor + \beta_{14} \cdot Furnished + \beta_{15} \cdot Terraced\ house + \beta_{16} \cdot House + u$$

5. RESULTS

This section is dedicated to give results of the estimated regressions.

6.1 GENERAL MODEL:

Table 4: Results of the general model

	(1)	(2)
Cons	4.568*** (70.03)	4.537*** (59.17)
Lsqm	0.602*** (32.79)	0.555*** (24.04)
Rooms	-0.024*** (-3.15)	-0.002 (-0.31)
Bathrooms	0.144*** (13.23)	0.127*** (9.45)

SOUTH	-0.654*** (-41.49)	-0.609*** (-43.82)
Terrace	-	-0.032*** (-2.79)
AC	-	0.146*** (12-67)
Garage	-	0.0007 (0.04)
Storage	-	-0.069*** (-4.07)
Pool	-	0.116*** (4.13)
Location	-	0.104*** (7.17)
Garden	-	-0.124*** (-5.04)
Elevator	-	0.057*** (4.01)
Floor	-	0.007*** (2.66)
Furnished	-	-0.11*** (-4.9)
Terraced House	-	-0.03 (-0.85)
House	-	-0.084 (-1.43)
R^2	0.673	0.706
F-Stat	(4, 3064) 1572.73	(16, 3052) 543.14

Note: Coefficients of the variables for the different regression with Lprice as dependent variable. T-statistic in braquets. (-) for variables not used in the model. All standard deviations are calculated as ***robust*** *because of previously checked heterocedasticity*** p<0.01. ** p<0.05. * p<0.1.*

In the general model (1) all variables are found to be statistically significant. The goodness of fit or R^2 gets the value of 67,3% which is a very acceptable value. It is estimated that the rent for a house with the same characteristics located in the **South** is **65,4% cheaper** than the ones in the City. The negative effect of **Rooms** is not in line with expectations and literature (see Table 1), an increase in 1 rooms leads to a decrease of 2.4% in the price of the rental house, ceteris paribus. **Bathrooms** and **square meters** present a positive sign as expected.

In order to better understand the negative sign of "Rooms", two other regressions with **interactions** have been tested. One with an interaction between South and Rooms and the second one with an interaction between City and Rooms. The conclusion derived from these regressions is that one more room decreases the rental price, specifically the South experiences a greater decrease in the rental price compared to one more room in

the Center, ceteris paribus. Complete results can be seen in Own elaboration based on (Ayuntamiento de Madrid, 2022)

APPENDIX B. Why could this happen? Suburbs are also called "Dormitory Cities" in which more families tend to live respective to the City centers where landlords are more attracted to compartmentalize the dwellings.

Talking about the results obtained in the model (2) the value of R^2 increases to 70,6% such that more than 70% of the rent price is explained by the independent variables included in the model. It must be said that the R^2 has only increased around 3% which lead us think the variables that mostly explain the price are the ones included in model (1).

Four of these variables are not significant in the model, these are "Rooms", "Garage", "Terraced House" and "House". The variables "Terrace", "Storage", "Garden" and "**Furnished**" are statistically significant but present a negative sign. This means, it is estimates that houses that are already furnished decrease the rental price an 11%, ceteris paribus. A possible explanation to why the variables "**Terrace**", "**Storage**" and "**Garden**" do not have the expected sign is that terraces in the city center are usually very small such that people prefer to have those few meters inside the house rather than having a small balcony or terrace. Storage rooms could be also affected in the same way. Having this in mind, it must be said that the decrease in the rental price when these three variables are present, is relatively small.

In relation with the variable **Elevator** one question came to mind, what is the relation between the floor level and having elevator? For this, model (2) was regressed again but creating an **interaction** between Floor and Elevator. (See Own elaboration based on (Ayuntamiento de Madrid, 2022)

APPENDIX B for full results). The main conclusion inferred from this interaction is that a higher floor level without elevator implies a lower price.

6.2 UNIVERSITY MODEL

6.2.1 MADRID-CITY UNIVERSITY AREA:

Table 5: Results of the Madrid-City university area model

	(3)	(4)
Cons	4.516*** (51.85)	4.516*** (54.83)
Lsqm	0.621*** (24.81)	0.557*** (22.46)
Rooms	-0.038*** (4.98)	-0.006 (-0.66)
Bathrooms	0.161*** (10.96)	0.133*** (8.74)

CITY UNIVERSITY AREA	-0.081*** (-6.05)	-0.067*** (-5.17)
Terrace	-	-0.035** (-2.53)
AC	-	0.161*** (11.78)
Garage	-	0.042** (1.70)
Storage	-	-0.085*** (-3.87)
Pool	-	0.195*** (4.62)
Location	-	0.0944*** (6.12)
Garden	-	-0.130*** (-3.25)
Elevator	-	0.066*** (4.01)
Floor	-	0.006** (2.04)
Furnished	-	-0.145*** (-4.62)
Terraced House	-	-0.007 (-0.17)
R2	0.622	0.663
F-Stat	(4, 2531) 976.61	(15, 2520) 332.84

Note: Coefficients of the variables for the different regression with Lprice as dependent variable. T-statistic in braquets. (-) for variables not used in the model. All standard deviations are calculated as ***ROBUST*** *because of previously checked heterocedasticity*** *p<0.01. ** p<0.05. * p<0.1.*

The City-Madrid University model shows interesting results, model (3) gets significant values for all variables. R^2 *is 6*2,2%. The study variable in this regression "City university area" shows a negative sign such that, it is estimated that a house rent located in the **City university area is 8,1% cheaper** than the exact same house located in a not university area in the City.

A negative relationship between Rooms and rent price appears as in the previous model, that is why a similar interaction (see Own elaboration based on (Ayuntamiento de Madrid, 2022)

APPENDIX B for complete results) between "number of rooms" and "City-university-area" has been created, getting a positive relationship between both. This leads us to conclude that, an additional room in the City university area decreases the price less than in the non-university City neighborhoods.

In model (4) R^2 increases by around 0,5% so the explanatory variables added in the model explain the 66,3% of the rent price.

There are two not significant variables, “Terraced House” and “Rooms”. The variables “**Storage**”, ”**Garden**” and “**Furnished**” present negative values meaning that when a house contains this commodities, the price of the rent decreases by β_x%. The same explanation as in the model (2) can be given. The coefficients with higher positive values are “**Air Conditioner**” and “**Bathrooms**” we can then say these commodities are very valued by the tenants.

6.2.2 SOUTH-MADRID UNIVERSITY AREA:

Finally, the South-Madrid university model shows different results to the ones in the City. In model (5) all variables are significant and explain the 69,1% of the variation in the rent price. The variable of interest in this model “South university area” shows the following results: it is estimated that a house rent located in the **South university area is about 12% more expensive** than the exact same house in the South not university area. The sign of the variable “**Rooms**” is now positive such that one more room in the house increases the rent price in 3,6%, ceteris paribus. Following the line from previous models, an **interaction** between “Rooms” and “South-university-area” has been done. Results suggest that the same house with one more room in a South University area increases more the price than one located in a non-university area of the South (See Own elaboration based on (Ayuntamiento de Madrid, 2022)

APPENDIX B) . Model (6) show that the variables “Terrace”, “Garage”, “Storage”, “Garden”; “Floor”, “Furnished”, “Terraced House” and “House” are not significant in this model and the variable “Pool” is significant but at a 10% significance level.

This results are interesting as this two models are the only ones having a positive coefficient for the variable “rooms” meaning is the only area which increases the rental price when increasing by 1 the number of rooms.

Table 6: Results of the South-Madrid University area

	(5)	(6)
Cons	5.186*** (62.69)	5.261*** (54.09)
Lsqm	0.277*** (12.18)	0.239*** (9.40)
Rooms	0.036*** (4.98)	0.045*** (5.89)
Bathrooms	0.14*** (9.17)	0.085*** (6.57)
SOUTH UNIVERSITY AREA	0.115*** (10.85)	0.119*** (11.61)
Terrace	-	0.009 (0.85)

AC	-	0.037*** (3.35)
Garage	-	-0.004 (-0.27)
Storage	-	0.009 (0.73)
Pool	-	0.030* (1.87)
Location	-	0.040* (1.77)
Garden	-	0.017 (1.07)
Elevator	-	0.0334** (2.37)
Floor	-	-0.001 (-0.28)
Furnished	-	-0.023 (-1.52)
Terraced House	-	-0.002 (-0.07)
House	-	0.0587 (1.63)
R^2	0.691	0.718
F-Stat	(4, 528) 275.42	(16, 516) 82.63

Note: Coefficients of the variables for the different regression with Lprice as dependent variable. T-statistic in braquets. (-) for variables not used in the model. All standard deviations are calculated as ***ROBUST*** *because of previously checked heterocedasticity*** p<0.01. ** p<0.05. * p<0.1.*

7 CONCLUSIONS

Housing is a necessary 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 (*Lancaster, 1966) (Rosen, 1974)* this has been a very used method to explain prices all around the world.

Given the post-pandemic era in which we are living, 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 attributes.

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 in which the dependent variable “Natural Logarithm of rental price” is explained by various characteristics such as “Number of rooms”, “Number of bathrooms” and other dummy’s such like “Air conditioning”, “Location” or “Pool”.

The results obtained lead to the following conclusions:

In general, the models adequate to the general literature. In relation with the comparison model between the City and South areas it is shown that the same house located in the **suburbs is 65% cheaper**. Bathrooms are very valued commodities as an increase in one bathroom leads to an increase of a 14% of the rental price. One of the variables that does not follow the current literature (Table 1) is Rooms. Thanks to the interaction made in the general model we can conclude that one more room leads to a decrease in the price, even though this decrease is greater in the South rather than in the City.

Between the **university models** that have been estimated, the most remarkable result is: The City university area is cheaper compared to the rest of the City whereas the South university area is more expensive than the rest of the South.

This study also emphasizes the importance of the control variables, “sqm”, “Bathrooms”, “Rooms” which in all models explain more than 60% of the final price.

A way of **improving this analysis** could be increasing the collection period. These 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 solutions. The variable “Year” has not been added to the models as it was not possible to find 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. As seen in previous literature using cross-sectional data is also interesting in this type of analysis as dynamics can be analyzed and are important in order to see how do prices in different areas react to business cycles.

8. 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 2008, *Introduction to Econometrics*, 4th edn, Cengage Learning.

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

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

18) 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.

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

20) 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.

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

22) 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.

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

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

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

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

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

28) 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.

29) 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.

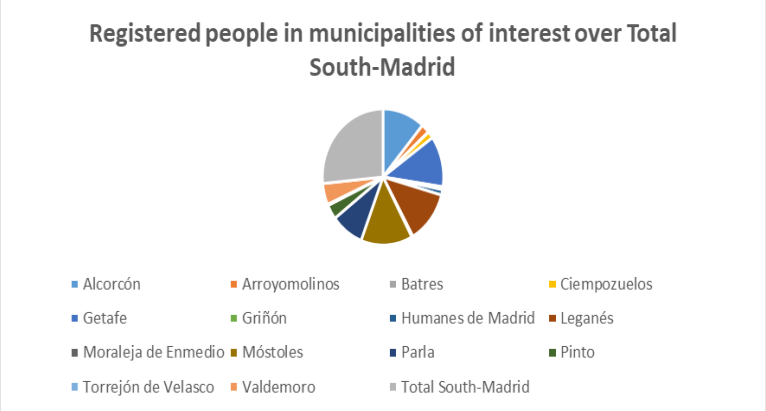
30) Wilson, B. & Frew, J. 2007, *Apartment Rents and Locations in Portland, Oregon*.

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

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

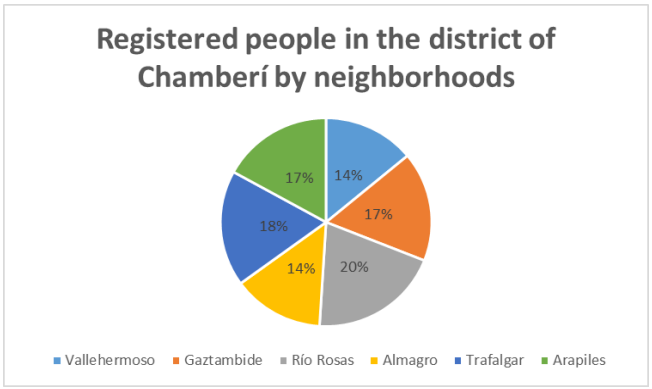
APPENDIX A

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



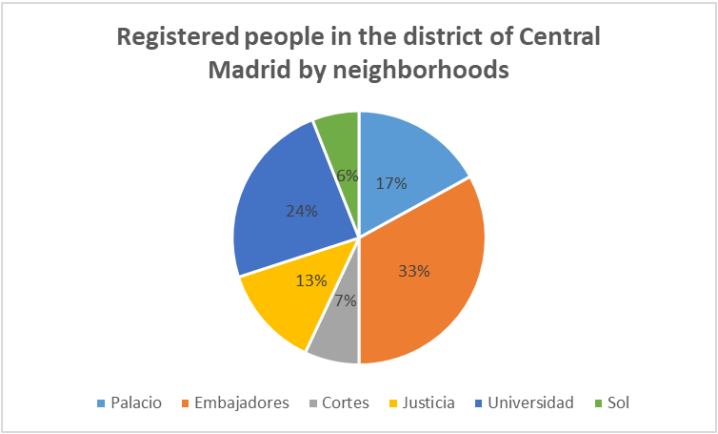
Own elaboration based on (Ayuntamiento de Madrid, 2022)

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



Own elaboration based on (Ayuntamiento de Madrid, 2022)

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



Own elaboration based on (Ayuntamiento de Madrid, 2022)

APPENDIX B

Table B-10: VIF and Residuals Plot for the General Model

VIF		
Variable	VIF	1/VIF
Lsqm	3.37	0.2967
Rooms	2.85	0.3513
Bathrooms	2.41	0.4141
Garden	1.96	0.5101
Pool	1.89	0.5278
Garage	1.66	0.6031
SOUTH	1.65	0.6045
Storage	1.33	0.7526
Location	1.2	0.8314
Terrace	1.17	0.8563
AC	1.13	0.8833
House	1.11	0.9026
Furnished	1.04	0.9636
Terraced House	1.01	0.9898

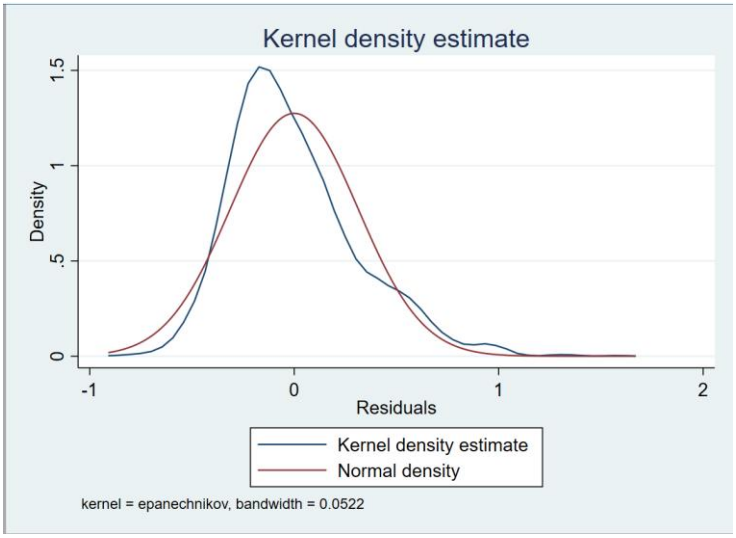


Table B-11: General Model interactions (ROOMS*SOUTH AND ROOMS*CITY)

INTERACTION MODEL (1)	Rooms*South	Rooms*City
Cons	4.572*** (57.46)	3.991*** (46.27)
Lsqm	0.5999*** (26.07)	0.5999*** (26.07)
Rooms	-0.0194*** (-2.24)	-0.050*** (-4.57)
Bathrooms	0.1425*** (11.09)	0.1425*** (11.09)
SOUTH	-0.5811 (-19.93)	-
CITY	-	0.5811*** (19.93)
Rooms*South	-0.0302** (-2.53)	-
Rooms*City	-	0.0302** (2.53)
R^2	0.673	0.673
F-Stat (5, 3063)	1690.17	1690.17

Table B-12: General Model interaction (Elevator*Floor)

INTERACTIONS MODEL (2)	ELEVATOR*FLOOR
Cons	4.608*** (56.96)
Lsqm	0.5502*** (23.74)
Rooms	-0.0007 (-0.10)
Bathrooms	0.1261*** (9.36)
SOUTH	-0.6115*** (-44.01)
Terrace	-0.032*** (-2.80)
AC	0.147*** (12.80)
Garage	0.001 (0.08)
Storage	-0.0684*** (-4.06)
Pool	0.117*** (4.16)
Location	0.105*** (7.25)
Garden	-0.125*** (-5.08)
Elevator	-0.0062 (-0.25)
Floor	-0.183*** (-2.12)
Furnished	-0.109*** (-4.90)
Terraced House	-0.0292 (-0.81)
House	-0.045 (-1.45)
Elevator * Floor	0.0281*** (3.07)
R^2	0.7074
F-Stat (17, 3051)	513.64

Table B-13: Vif and residuals plot for City-University area model

VIF		
Variable	VIF	1/VIF
Lsqm	3.55	0.2818
Rooms	2.9	0.3443
Bathrooms	2.5	0.4002
Garden	1.39	0.7173
Pool	1.28	0.7835
Garage	1.26	0.7923
Location	1.2	0.8323
Terrace	1.17	0.8517
Storage	1.17	0.8545
AC	1.12	0.8893
Elevator	1.1	0.9061
CITY UNIVERSITY AREA	1.07	0.9317
Floor	1.06	0.9404
Furnished	1.03	0.9705
Terraced House	1.01	0.9867

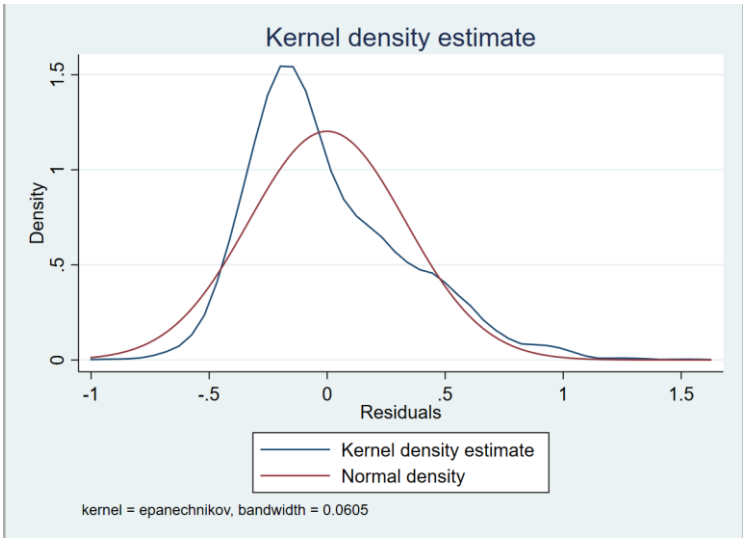


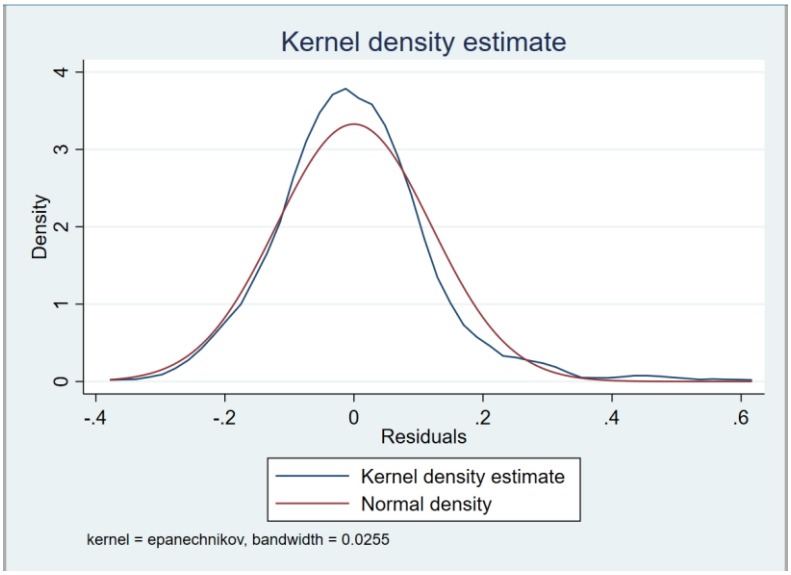
Table B-14: City-University-Area interaction (City-uni-area*Rooms and No-City-uni-area * Rooms)

INTERACTIONS MODEL (3)	CITY UNIVERSITY AREA * ROOMS	NO-UNIVERSITY AREA CITY * ROOMS
Cons	4.535*** (52.38)	4.411*** (52.15)
Lsqm	0.62*** (25.27)	0.62*** (25.27)
Rooms	-0.046*** (-4.55)	-0.022** (-1.96)
Bathrooms	0.163*** (11.44)	0.163*** (11.44)

NO-UNIVERSITY AREA CITY	-	-0.122*** (4.90)
CITY UNIVERSITY AREA	-0.122*** (-4.90)	-
NO-UNIVERSITY AREA CITY * ROOMS	-	0.025*** (-2.25)
CITY UNIVERSITY AREA * ROOMS	0.025*** (2.25)	-
R2	0.623	0.623
F-Stat (5, 2530)	808.74	808.74

Table B-15: VIF and Residuals Plot for South-University-area model

VIF		
Variable	VIF	1/VIF
Lsqm	3.59	0.2784
Rooms	2.32	0.4309
Bathrooms	2.31	0.4337
Garden	2.26	0.4433
Pool	2.2	0.4538
Garage	1.8	0.5555
House	1.7	0.5867
Storage	1.54	0.6501
Elevator	1.46	0.6837
AC	1.2	0.8355
Floor	1.18	0.8469
SOUTH UNIVERSITY AREA	1.12	0.8919
Terrace	1.13	0.8876
Terraced House	1.08	0.9271



Furnished	1.08	0.9287
Location	1.05	0.9502

Table B-16: South-University-Area interaction (South-uni-area*Rooms and No-South-uni-area * Rooms)

INTERACTIONS MODEL (3)	SOUTH UNIVERSITY AREA * ROOMS	NO-UNIVERSITY AREA SOUTH * ROOMS
Cons	5.204*** (62.43)	5.266*** (61.59)
Lsqm	0.281*** (12.21)	0.281*** (12.21)
Rooms	0.219** (2.87)	0.431*** (5.03)
Bathrooms	0.105*** (9.23)	0.105*** (9.23)
NO-UNIVERSITY SOUTH AREA	-	-0.063** (-1.97)
SOUTH UNIVERSITY AREA	-0.0629** (1.97)	-
NO-UNIVERSITY SOUTH AREA * ROOMS	-	-0.0212* (-1.73)
SOUTH UNIVERSITY AREA * ROOMS	0.0212* (1.73)	-
R2	0.692	0.692
F-Stat (5, 527)	228.30	228.30