Universidad Autónoma de Madrid

Madrid, 13 de junio de 2024

Estimados miembros del comité organizador,

Confirmo que Marina Gómez García es estudiante de doctorado de segundo año a tiempo parcial en el programa de Economía y Empresa de la Universidad Autónoma de Madrid. Marina está bajo la supervisión conjunta de Omar Rachedi de ESADE y de mí, Zoë Kuehn, profesora de la Universidad Autónoma de Madrid.

Quedo a su disposición para cualquier información adicional que puedan requerir.

Atentamente,

Zoë Kuehn
Associate Professor,
Universidad Autónoma de Madrid
E-mail: zoe.kuehn@uam.es
Web: https://sites.google.com/site/zoekuehn

The impact of air pollution (NO_2) on the real estate market: the case of Madrid city [1]

Author: Marina Gómez-García [a,*]

[a] **Autónoma of Madrid University and Banco de España**

Abstract:

Air pollution has been found to have widespread negative effects for individuals' health and labor market outcomes, even determining migration decisions. Hence, it is natural to think that air pollution might also affect the value of location-specific assets such as housing. I analyze how air pollution affects housing prices in Madrid city, a large European capital with high levels of pollution and high property rates. Using location-specific information on air pollution and the universe of housing transactions, I estimate the causal impact of air pollution on housing prices, controlling for house and neighborhood characteristics, time, and neighborhood fixed effects. To address endogeneity concern, I exploit quasi-experimental variation in nitrogen dioxide levels resulting from thermal inversions. The results suggest that an increase of 1 % in air pollution reduces housing prices by 0.058 %, and this effect is non-linear, being larger for higher levels of pollution. Furthermore, I find that high income or more educated neighborhoods, as well as those with a higher share of foreign-born residents, are willing to pay more for clean air. These findings are robust to alternative measures of air pollution, interpolation techniques, and the inclusion of additional controls such as traffic noise.

[1] The author would like to thank her supervisors, Professor Zoe Kuehn and Omar Rachedi for their valuable comments and suggestions. All remaining errors are my own. The ideas in this paper are my own and do not necessarily represent the views of Banco de España.

* Corresponding author.

Email address m.gomez@bde.es (M. Gómez-García)

1.Introduction

Air pollution has important impacts for individuals' health, subsequently affecting economic outcomes such as work productivity.[2] As a result, national and international institutions have implemented more stringent regulations on air pollution in recent decades. Furthermore, numerous surveys across Europe and Spain agree that people are aware and conscious about the negative consequences of air pollution, and there is a positive correlation between the level of air pollution and the level of concern (European Commission Eurobarometer (2019, 2022); Dons et al., 2018). Forsberg et al. (1996) conduct a survey in Sweden regarding the perception of air quality and results suggest that nitrogen dioxide (NO_2) concentrations correlate consistently with reported annoyance related to air pollution. There is also evidence that people are not only aware of air pollution but that it also affects important individual decisions such as where to live. For instance, Mikula & Pytlikova (2021) find a 24% reduction in emigration after improvements in air pollution in the Czech Republic, as better technology was implemented in coal- burning power plants (without affecting economic activity). Chen et al. (2022) find that a 10 percent increase in air pollution reduces local population by 2.8 percent in China, using thermal inversion as an instrumental variable for air pollution.

Given that air pollution has a widespread range of negative effects and individuals are aware of that, it is natural to think that air pollution can affect asset values which are location specific, such as housing. The aim of this paper is to provide a causal estimate of how air pollution concentrations affect housing prices, focusing on Madrid, a big European city with high levels of pollution and high property rates. Indeed, Madrid has the highest mortality burden attributable to NO_2 pollutants (Lancet, 2021), and according to EU-SILC (2018) property rates in Spain are around 76% and are much higher than in other European countries such as France (65%), Germany (52%), the United Kingdom (65%), and Italy (72%).

My empirical analysis is based on Land Registry data which includes the exact location of each transaction, and I estimate the impact of air pollution on housing prices across neighborhoods and years. My baseline model is a linear regression model of housing prices as a function of local air pollution over the three months prior to the transaction, controlled by time-fixed effects (years), time-invariant neighborhood fixed effects as well as a set of time-varying neighborhood characteristics. However, this last set of controls is limited and hence there might still be omitted variables bias leading to endogeneity issues. Hence, I apply an instrumental variable regression approach exploiting the variation in NO_2 pollutants driven by thermal inversions. Thermal inversions are an atmospheric phenomenon that causes a deterioration in air quality because the air cannot move freely and remains trapped near the surface. My estimates suggest that a 1% reduction in air pollution leads to a 0.058% decrease in housing prices. These results are robust to alternative measures of air pollution, interpolation techniques, and the inclusion of additional controls such as traffic.

The current paper provides new evidence on the causal effects of air pollution on housing prices for a developed country in Europe using a natural phenomenon as an instrument. Most previous studies, on the other hand, use environmental regulation as an instrument variable to investigate the effects of pollution on housing prices. For instance, Gruhl et al. (2022), observe

[2] For instance, Chay & Greenstone, 2003; Guarnieri & Balmes, 2014 show how air pollution increases the probability of asthma, Cohen & Pope, 1995; Fredrik et al., 2000 focus on the effect on lung cancer and Chay & Greenstone, 2003; Deryugina et al., 2019 on mortality. Hanna & Oliva, 2015; Holub, Hospido & Wagner, 2021 find increases in work productivity when air quality improves.

an increase in apartment rents and property prices upon the introduction of low-emission zones in Germany, while C. K. Tang (2016) study the evolution of housing prices in London after the introduction of a congestion charge that improved traffic conditions. Although in both cases housing prices increase because of the change in regulation, it is difficult to distinguish improvements in air pollution from reductions in traffic and traffic noise. Other existing studies which use linear regression model to estimate the effect of air pollution on housing prices, are Amani et al. (2022) who find that an increase of 10% in air pollution decreases housing prices in 0.6-0.8% while Le Boennec and Salladarré (2017) find that air pollution does not affect housing prices in Nantes, a city with air pollution well below established limits.

Closely related is a recent paper by Cai et al. (2024) who document the relationship between housing prices and air pollution in Beijing also using thermal inversion as instrumental variable. My paper differs from Cai et al. (2024) as I consider that buying a house is a time-intensive process whereas they only consider average quarterly in $PM_{2.5}$. In fact, 50% of buyer takes at least 4 months in 2017 to buy a house in Spain (Fotocasa research, 2017). Therefore, buyers are likely to consider pollution levels not only for the deed month but also for the preceding three, six or nine months when making a purchasing decision. Additionally, I applied the same criteria for other neighborhood variables, and I conducted regression analyses for the preceding three, six and nine months. The results suggest that individuals are more likely to base their house-buying decisions on longer-term air quality readings.

Another aim of this paper is to study the heterogeneous effects of air pollution looking to different levels of pollution (Chay & Greenstone, 2005; Amani et al., 2022). This paper shows that housing prices are not affected if pollution is under the limits by looking at non-linearities of air pollution. Finally, this study provide evidence on the heterogeneous effects of air pollution on high educated and high-income neighborhoods (Cai et al., 2024; Mikula and Pytlikova, 2021), also contribute to the literature on migration as results suggest that neighborhoods with high percentage of foreign people are willing to pay more for clean air.

The remainder of this paper is organized as follows: The next Section presents the empirical model, and in Section 3 I present the data. Section 4 discusses the results and provides robustness checks, and Section 5 concludes.

2. Methodology

2.1 Fixed effect model

To analyze the effect of air pollution on housing prices in Madrid, I run the following neighborhood and time fixed effects estimation:

$$\ln P_{igdt} = \beta_0 + \beta_1 \frac{1}{3}\sum_{t}^{t-2} \ln NO_{2\,gdt} + \delta_2 X_{igdt} + \delta \frac{1}{3}\sum_{t-2}^{t} N_{gdt} + \alpha_g + \alpha_y + \alpha_{dy} + w_{igdt}, \quad (1)$$

where the dependent variable ($\ln P_{igdt}$) is the logarithm of housing price per square meter, indices i, g, d and t refer to the transaction, neighborhood, district and month of transaction, respectively. $\frac{1}{3}\sum_{t}^{t-2} \ln NO_{2\,gdt}$ is the logarithm mean of air pollution over the last three months. X_{igdt} are housing characteristics such as property type (new or second-hand), property size, and year of construction, and N_{gdt} are the mean of neighborhood characteristics such as the number of businesses activities and population over the last three months. α_g and α_y are neighborhood fixed effects to capture time-invariant determinants of housing prices in a neighborhood and year fixed effects to control for time-varying factors which affect housing prices across Madrid in different years. To address concerns of increasing housing prices due to the proliferation of short-term rentals like Airbnb, I also control for an interaction term between year and the group of districts (α_{dy}) where this type of rental is more common.[3] Finally, ε_{igdt} is the idiosyncratic error term and the standard errors are clustered at the neighborhood level.

My coefficient of interest is β_1 which reflects the relationship between air-pollution and housing prices within a neighborhood, and I expect it to be negative and significant. While equation (1) attempts to address endogeneity concerns controlling for neighborhood-fixed effects, business activities, population density and year-fixed effects, it is challenging to account for all variables that may influence housing prices and air pollution. For instance, improvements in public transportation, such as extended bus schedules or new bus stops along a neighborhood route, could potentially increase housing prices and reduce air pollution. Conversely, certain areas with high daytime population density due to work-related activities may experience increased pollution levels. However, despite this pollution, individuals may still prefer to live in such areas to minimize commuting time, which can lead to both higher air pollution levels and housing prices.

2.2 Fixed effects two-stage least squares (FE2sls)

I address the endogeneity concern present in equation (1) using a two-stage least squares (2SLS) framework. This approach uses quasi-experimental variations in NO_2 levels attributed to thermal inversion phenomena.

I denote by $\frac{1}{3}\sum_{t}^{t-2} \ln Z_{gdt}$, the logarithm of the mean of thermal inversion in neighborhood g, district d and month t, over the last three months. The first stage IV equation is:

$$\frac{1}{3}\sum_{t}^{t-2} \ln NO2_{gdt} = \beta_0 + \beta_1 \frac{1}{3}\sum_{t}^{t-2} \ln Z_{gdt} + \delta_2 X_{igdt} + \delta_3 \frac{1}{3}\sum_{t}^{t-2} N_{gdt} + \alpha_g + \alpha_y + \alpha_{dy} + \varepsilon_{igdt,} \qquad (2)$$

where all variables except $\frac{1}{3}\sum_{t}^{t-2} \ln Z_{gdt}$ correspond to those in equation (1). The second stage equation estimates housing prices after substituting predicted pollution $\ln NO_2{}^*_{gdt}$ from equation (2) for $NO_{2\,gdt}$ in equation (1)

$$\ln P_{igdt} = \beta_0 + \beta_1 \frac{1}{3}\sum_{t}^{t-2} \ln NO2^*_{gdt} + \delta_2 X_{igdt} + \delta_3 \frac{1}{3}\sum_{t}^{t-2} N_{gdt} + \alpha_g + \alpha_y + \alpha_{dy} + w_{igdt}, \qquad (3)$$

where $\ln NO_{2\,gdt}$ is the predicted ambient NO_2 concentration in each neighborhood g, district d and month t.

Two key assumptions for this approach are that thermal inversion has no direct effect on housing prices other than via increased ambient concentrations of air pollution, and that thermal inversion is correlated with air pollution.

3. Data

My data set combines information on housing prices and air pollution for Madrid city for 2014-2020 which I complement with additional information on neighborhood characteristics and data on thermal inversion. Madrid, Spain's capital, has a population of around 3.2 million

[3] Central, Salamanca, Chamberí, Arganzuela and Retiro districts.

people, representing 7% of Spain´s total population. There are 131 neighborhoods in Madrid which are grouped in twenty-one different districts. Figure A1 in the Appendix displays the distribution of neighborhoods and districts across Madrid city.[4] Note that districts have certain autonomy regarding municipal affairs (for instance they can process and solve applications for urban planning licenses or maintain park and garden facilities and services assigned to the district).[5] However, I observe quite variation across neighborhoods even within one district which is why I prefer to carry out the analysis at the neighborhood instead of the district or zip-code level.[6]

Housing Prices

Data on housing prices comes from Land Registry records which encompass all property transactions made in Madrid city. This dataset includes detail information on transaction price and date, built square meters, location, property type (detached, semi-detached, flat with and without buildings) and whether the property is new or second-hand. If there is missing data in one variable used, I eliminate the transaction. Also, if I observe duplicates of transactions with same date or same cadastral reference with differences across certain variables of interests, such as price or square meter, I drop both transactions. Since the year of construction which might have an important impact on housing prices is not available in the Land Registry data, I supplemented it with cadastral data from Moral- Carcedo (2023) that includes this information.

All transactions are geocoded using cadastral references, which allows me to assign each transaction to a specific neighborhood. I calculate the actual price per square meter, using Consumer Price Index (CPI) from the National Statistics Institute, with a reference date of February 2020.

Air pollution

Daily or hourly data on air pollution for Madrid City are available from the open data of the Madrid City Council website. In Madrid, there are 24 monitors that measure the concentration of NO_2 (Figure A2 in appendix): 9 are traffic monitors close to a main road (red points), 12 monitors are urban background stations whose pollutant levels are representative of the average exposure of the general population (blue points), and 3 suburban monitors which are located on the outskirts of the city (green points).[7] To compute measured air pollution using NO_2, I consider the monthly median for each monitor.[8] Finally, I compute the mean over the last three, six and nine months which I use as the value of air pollution.

---

[4] The number of neighborhoods was not constant during the study period, three new neighborhoods were created: (i) Ensanche de Vallecas from district 18 (Villa de Vallecas), (ii) El Cañaveral, and (iii) Valderribas from district 19 (Vicálvaro) in 2017. Although the number of districts is constant during the study year, as a robustness check, I eliminate districts 18 and 19 to ensure that results do not change.

[5] See Boletin Oficial Ayuntamiento de Madrid (2023) or https://sede.madrid.es/csvfiles/UnidadesDescentralizadas/UDCBOAM/Contenidos/Boletin/2023/Julio/Ficheros%20PDF/BOAM_9421_06072023141357838.pdf for details.

[6] Within the Salamanca district, the neighborhood Recoletos which is close to Retiro park and to Castellena street has an average household income of 80.000 euros whereas it is 42.000 euros in the Fuente del Berro neighborhood. Zip codes do not follow any administrative rules, so one zip code can aggregate neighborhoods from different districts.

[7] All monitors remain in the same place during the dates of analysis (Jan 2014 - Feb 2020), and there is data available for all months.

[8] I first compute the daily mean for each monitor, and then conditional on having at least 9 observations for each month I compute the monthly median of each monitor. All monitors satisfy this condition, so

To assign air pollution levels to each neighborhood, I use the so-called "Kriging" interpolation method (Anselin & Le Gallo, 2006). Kriging is a spatial interpolation technique that estimates values for unsampled locations based on nearby sampled points, considering spatial correlation and variability. By applying Kriging, I can generate a continuous surface of pollutant concentrations across the area of study, allowing me to assign pollution levels to specific neighborhoods based on their geographic locations. For robustness I also consider an alternative methodology which consists of an inverse distance-weighted average of the three closest stations to each neighborhood (Amani et al., 2022).

Madrid Central is a permanent measure to reduce air pollution in the center of Madrid city. It was announced on 30th of October 2018 and its implementation was on 30th of November 2018, but it was not until 15th of March 2019 that it started to issue fines. The aim of this measure was the creation of a special area in the central district (district 1) where only the cars of people living there and eco or zero-emission labeled cars can drive or park in that area. [9] As Madrid Central is less than 1% of Madrid city and its introduction was at the end of my analysis it doesn't affect my results. [10]

Thermal inversions

Thermal inversion is a meteorological phenomenon. Whenever it occurs the normal decrease in air temperature with increasing altitude is inverted, resulting in a layer of warm air forming above a layer of cooler air near the ground (see Figure A3). This can cause pollutants and moisture to become trapped near the ground leading to an increase in air pollution. The data used for analyzing thermal inversion comes from Climate Data Store on the Copernicus Website.[11] This dataset provides hourly air temperature data across 37 vertical layers from 1940 to 2023, with a spatial resolution of approximately 0.25°x0.25° (about every 25km x 25km). Thermal inversion is then calculated in the following way: First, I create a binary variable equal to one if the hourly temperature difference between atmospheric layers 900 and 875 is negative, and zero otherwise. I then compute the monthly mean of this dummy variable and I compute the mean over the last three months which I use as a proxy for thermal inversion conditions for each month. As in the case of air pollution, Kriging interpolation is applied to assign thermal inversion to each neighborhood.

Neighborhood data

I obtain data on number of shops, restaurants, educational institutions (schools, day care centers, language academies or driving school), entertainment and recreation centers and other services at neighborhood and month level using the Registry of establishments, activities,

---

there is observation for all months and monitors. Note that I use NO_2 because it is one of the main pollutants in the national index and there are 24 air pollution monitors that measure this pollutant for further details see Appendix B.

[9] In general terms, electric cars and plug-in hybrid that cover at least 40km are zero emission labeled. Those cars plug-in hybrid or hybrid or powered by natural gas will be Eco- label. C labels are those with EURO 4/IV, 5/V or 6/VI gasoline and EURO 6/VI diesel. B label are EURO 3/III gasoline and diesel Euro 4/IV or 5/V and those not including in these categories will be A label that is the same as without label and are the most polluting vehicles.

[10] As a robustness check, district 1 is eliminated from the sample in section 5.4.1.

[11] This website supports society by providing authoritative information about the past, present and future climate in Europe and the rest of the World (see https://climate.copernicus.eu/about-us).

and hospitality and restaurant terraces available on Madrid city's open data portal.[12] I compute the number of services (education, health, wholesale and retail, accommodation, entertainment and recreation and other services) for each month and neighborhood and after that I compute the mean over the last 3 months.

In addition, total population in each neighborhood and month is an important variable that I obtain from the census of Madrid city available on their webpage. Also, this information has been computed as a mean over the last 3 months. Finally, data on Airbnb rentals is obtained from the Airbnb webpage, using data for September 2022, the closest available data. Gil & Sequera (2020) show that Airbnb rentals in Madrid have increased significantly since 2015, and they are concentrated in the city´s central district (District 1). This district represents less than 1% of the surface of Madrid city. Table A1 in the Appendix displays Airbnb rentals in each district by household, I select the five neighborhoods with the highest concentration of Airbnb rentals by households and control for potential different time trends in housing prices within this group of neighborhoods.

I also use data from a Household Panel carried out by the city council of Madrid which is available on the webpage of the city council of Madrid from 2017 onwards. [13] This data includes information on average household income, education and share of foreigners for each neighborhood. This information is useful for the heterogeneity analysis made in section 5.4. As data is yearly and only available since 2017, I used for this analysis the information refers to 2017 which is the middle year of my study.

<u>Traffic data</u>

Finally, as traffic is very related to air pollution, I will control for traffic as an additional robustness check. The City Council of Madrid provides hourly traffic data on their website since 2013, I use the "occupation" variable which represents the percentage of time a traffic detector is active. For instance, an occupation of 50% means that over a period of 15 minutes, vehicles have been located over the detector for 7 minutes and 30 seconds. Using the specific location of the traffic monitors, each monitor is linked to the closest neighborhood based on the centroid coordinates of each neighborhood.

<u>Sample</u>

Sample data started in 2014, as information on neighborhoods characteristics is only available from that year. Since COVID-19 began in March of 2020, and Spain experienced a strict lockdown where only essential activities were permitted and free movement was restricted, buying a house was almost impossible and air pollution also dropped drastically. To avoid noise in this analysis, I conclude it in February 2020.

Additionally, I only select dwellings purchased by a natural person whose prices are set in by the open market and involving complete property acquisitions. This approach focuses on

---

[12] Retrieved from https://datos.madrid.es/portal/site/egob/menuitem.c05c1f754a33a9fbe4b2e4b284f1a5a0/?vgnextoid=23160329ff639410VgnVCM2000000c205a0aRCRD&vgnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnextfmt=default

[13] Retrieved from https://datos.madrid.es/portal/site/egob/menuitem.c05c1f754a33a9fbe4b2e4b284f1a5a0/?vgnextoid=71359583a773a510VgnVCM2000001f4a900aRCRD&vgnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnextfmt=default

individuals who might be concerned about their health when buying a house, thus excluding legal entities and acquisitions through inheritance or donation. The final sample has 105.320 observations of housing transactions between January 2014 and February 2020 in Madrid city.

Moreover, Figure A2 shows the location of monitors, the three suburban monitors are situated within large parks such as El Pardo, Casa de Campo and Barajas. To minimize potential distortions, I have excluded data from these suburban monitors in the main analysis. Additionally, I have omitted information from the monitor situated inside Retiro Park due to its controversial location and for the same reason as the suburban monitors. As robustness check, I will conduct additional analyses that include data from these excluded monitors to assess their impact and validate the main findings.

Descriptive statistics

Table 1 shows the descriptive statistics of main variables. The actual price per square meter is 2,723 euros from 2014 to 2020 and most of the houses are flats without buildings (do not have garage or storage room). Most dwellings that were bought and sold between 2014 and 2020 in Madrid city had been constructed before 1975, whose mean size is around 75 m2 and are second-hand. Figure A4 in appendix show the annual variation of housing prices across neighborhoods between 2016 and 2017, the difference on price is almost cero for El Pardo in the north and Butarque in the south, while Castillejos in Tetuan district has the biggest increase in housing prices.

The NO_2 concentration is reported in Panel 2, Table 1. The mean of NO_2 between 2014 and 2020 is between 38 $\mu g/m^3$ and 39 $\mu g/m^3$, which is quite close to the annual limit stablished by European Union, which is 40 $\mu g/m^3$. For some monitors its maximum has exceeded 70 $\mu g/m^3$ in the last three months, 64 $\mu g/m^3$ and 58 $\mu g/m^3$ in the last six and nine months, respectively. Figure A5 in appendix shows the variation of air pollution during 2016 and 2017 by neighborhoods. The south, Cuatro Vientos in the west and Villaverde district, are areas more polluted than El Pardo and the north areas.

Figures A6 in the Appendix show the variation between 2016 and 2017 in thermal inversions. Comparing this figure with Figure A5 (air pollution), it shows that there is a high correlation between air pollution and thermal inversion across neighborhood. But there are some differences when comparing these figures, for example Casa de Campo and Barajas neighborhoods.

Finally, regarding other neighborhood characteristics, the average neighborhood includes 30,800 individuals, ranging from 1,026 in Atocha neighborhood to 66,000 in Vista Alegre neighborhood. Most widespread activity is wholesale and retail stores, followed by accommodation and other services.

Table 1: Descriptive statistic

Variable	Mean	Sd	Min	Max	N
1. Land Registry data	2,723.07	1,548.23	136.84	9,931.94	105,320
Price/m^2	7.74	0.61	4.92	9.2	105,320
Less than 1960	0.28	0.45	0	1	105,320
Between 1960 - 1975	0.36	0.48	0	1	105,320
Between 1975 - 2000	0.19	0.39	0	1	105,320
Greater than 2000	0.17	0.38	0	1	105,320
Flat without buildings	0.84	0.37	0	1	105,320
Flat with buildings	0.14	0.34	0	1	105,320
Semi-detached	0.01	0.12	0	1	105,320
Detached	0.01	0.09	0	1	105,320
Less than 55 m^2	0.24	0.43	0	1	105,320
Between 55 and 75 m^2	0.26	0.44	0	1	105,320
Between 75 and 100 m^2	0.29	0.45	0	1	105,320
Greater than 100 m^2	0.21	0.4	0	1	105,320
Second-hand	0.83	0.37	0	1	105,320
2. NO_2 concentration (µg/m^3)					
Last 3 months	38.39	10.11	17.18	69.88	105,320
Last 6 months	38.54	7.69	18.24	63.88	105,320
Last 9 months	38.56	5.49	19.64	58.03	105,320
3. Thermal inversion					
Last 3 months	0.05	0.05	0	0.24	105,320
Last 6 months	0.05	0.03	0	0.13	105,320
Last 9 months	0.04	0.02	0.01	0.1	105,320
4. Neighborhood characteristics					
Total population	30,821.58	13,657.18	1,026.00	66,571.00	105,320
Mean of human health	40.21	19.92	1	113.33	105,320
Mean education	42.27	21.53	1	127.33	105,320
Mean wholesale and retail	496.88	338.87	2.67	1705.33	105,320
Mean accommodation	190.84	127.51	1.33	672	105,320
Mean entertainment and recreation	27.65	15.16	1	75	105,320
Mean other services	118.97	62.03	1	294.67	105,320

5. Results

5.1 Benchmark regression results

Table 2 reports OLS estimates (in column 1), IV estimates (column 2) and the first stage (column 3). All regressions control for house and neighborhood characteristics and year fixed effects as well as neighborhood fixed effects. Results from the OLS estimation of equation (1) suggests that there is a relationship between housing prices and NO_2 of -0.047%; meaning that an increase of 1% in air pollution will decrease housing prices by 0.047%.

The IV regression is implemented in a two stage least squares (2SLS) procedure where NO_2 is first regressed on the thermal inversion variable and all controls (see column 3), and in the second stage the log of housing price per square meter is regressed on the predicted NO_2 variable and all controls (column 2).

The causal effect of air pollution on housing prices is -0.058%, suggesting that an increase of 1% in air pollution will decrease the house prices by 0.058%. This represents a stronger negative effect than the one obtained via OLS meaning that omitted variables seem to mitigate the impact of air pollution on housing prices. Note that my estimates are in line with those by Cai et al. (2024) for Bejing where households are willing to pay an extra 0.0852% per housing unit price for an average quarterly reduction in PM2.5 of 1 µg/m3. Also, Amani et al. (2022) find that a 1% increase in NO_2 in the last three months lead to a decrease in housing of 0.061%, which is in line with my results.

Table 2: Baseline estimates for NO_2

	(1)	(2)	(3)
	House price per square meter		Log NO_2
	-0.047***	-0.058***	
Log NO_2	(0.000)	(0.000)	
			0.21***
Thermal inversion			(0.000)
	0		
Estimator	OLS	IV	First-stage
Observations	105,320	105,320	105,320
R-squared	0.607	0.607	0.75
First-stage F statistic		110.8	
House characteristics	YES	YES	YES
Neighborhood	YES	YES	YES
Year	YES	YES	YES
Union Airbnb* year	YES	YES	YES

The dependent variable for columns (1) and (2) is the logarithm of housing price per square meter and for column (3) is the logarithm of NO2 in the last 3 months. All columns include house characteristics and neighborhood dummies, year and neighborhood fixed effects and group of districts times year fixed effects. Robust standard errors in parentheses are clustered by year. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A2 and A3 in the Appendix present the full set of coefficients for OLS and IV regressions respectively including variables sequentially. Both estimation methods show similar results in coefficients and significance. Regarding coefficients, results are as expected once I control by neighborhood characteristics and neighborhood fixed effects (column 3). In column (4) I include time fixed effects and in column (5) I include the possibility of different time trends for neighborhoods with a high number of Airbnb, which is my baseline model. In both estimations, as the older is the building, higher is the drop in housing prices. A similar pattern occurs with property size, as the property is bigger, the individuals are willing to pay less per square meter. Finally, the number wholesale and retail, accommodation, and entertainment do not seem to influence a lot on housing prices.

5.2 Increasing Time Horizon.

One concern, as mentioned before, is that buying a house takes time, which makes it unlikely to observe contemporaneous effects of changes in air pollution levels on housing prices. This is why in the baseline model, I compute the mean of air pollution levels over the three months prior to the transaction, but buying a house might take longer, four months in median, (Fotocasa research, 2017). Therefore, I increase the time horizon and compute the mean of air pollution levels over the last six and nine months. I also increase the time horizon for other neighborhood characteristics and thermal inversion. Results in Table 3 seem to suggest that individuals are likely to take into account a longer- term air quality reading when buying a house. An increase of 1 % of NO_2 in the last three, six or nine months will decrease housing prices by 0.058%, 0.1% and 0.145% respectively. This also implies that in areas where pollution is persistent, buyers will offer less money for the purchase of a house, which is consistent with findings by Amani et al. (2022) for Teheran.

Table 3: Regression results increasing time horizon

	(1) Last 3 months	(2) Last 6 months	(3) Last 9 months
Log NO2	-0.058***	-0.100***	-0.145***
	(0.000)	(0.000)	(0.000)
Observations	105,320	102,052	99,191
R-squared	0.607	0.607	0.606
House characteristics	YES	YES	YES
Neighborhood	YES	YES	YES
Year	YES	YES	YES
Union Airbnb* year	YES	YES	YES

The dependent variable is the logarithm of housing price per square meter. The coefficient of interest show how housing prices change when considering air pollution over the last 3 months (column 1), the last 6 months (column2) and the last 9 months (column 3). All columns include house characteristics and neighborhood dummies, year and neighborhood fixed effects and group of districts times year fixed effects. Robust standard errors in parentheses are clustered by year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.3 Robustness checks

I check the robustness of my results, addressing aspects such as different housing price trends in certain districts, different measures of air pollution and interpolation methods or including additional control such as traffic occupation.

5.3.1 Eliminating Madrid's central district or districts with new neighborhoods

Two main concerns affecting housing prices in Madrid city were related to (i) Madrid Central and (ii) Airbnb. These concerns specifically affect the Central district (District 1) whose housing market has been strongly affected by tourism since 2015 and by traffic restrictions since 2019. Madrid's central district is made up of six neighborhoods: Palacio, Embajadores, Cortes, Justicia, Universidad, and Sol. To ensure the robustness of my findings, I exclude this district from the analysis. The results, summarized in Panel 1 of Table A4, indicate that although the estimates are slightly smaller, they align with my main findings.

In the second panel of Table A4, I exclude districts 18 (Villa de Vallecas) and 19 (Vicálvaro) due to the creation of new neighborhoods in these districts. This restructuring could influence housing price trends in these areas due to an increase in population and transactions. However, results suggest that the outcomes are robust and similar to my main results.

5.3.2 Traffic

Air pollution is very related to traffic, indeed, around 70% of NO_2 emissions are generated by traffic. I add my measure of traffic ("occupation variable") to the regression. Results in Table A5 suggest that when controlling air pollution already there is no additional effect of traffic on housing prices at the neighborhood level and coefficients for air pollution are indistinguishable from those in my baseline model. I tried other traffic measures such as intensity which is the number of vehicles per hour or track load which is an algorithm based on occupation and intensity that goes from 0 (empty) to 100 (collapse), and I found similar results. These results make sense as this analysis is made at neighborhood level, so I compute the mean of traffic congestion per neighborhood. There are neighborhoods with no information on traffic that will be dropped from the analysis.

5.3.3 Alternative air pollution measures and interpolation techniques

To check that my results are not driven by the interpolation technique used or by the measure of air pollution selected. Panel 1 of Table A6 shows the results for air pollution mean and the results are slightly bigger as the monthly mean is higher than the monthly median, but results show same tendency when I increase time horizon. To check other interpolation techniques, I follow Amani et al. (2022) that create an interpolation measure for pollution based on the inverse distance weighting technique. This technique consists of calculating the distance between the centroid of the neighborhood and air pollutant monitor, then I calculate the inverse distance- weighted average of the three closest monitors for each neighborhood and month. The results are shown in Panel 2 of Table A6 and the estimates are quite like the ones obtained by Kriging interpolation.

5.3.4 Including monitors from suburban areas and parks ("Retiro")

In Table A7, I add those monitors that I exclude from the main analysis. In the first panel, I add Retiro monitor that I eliminate it because it is inside Retiro park and its pollution levels can be interfered by trees and could not be representative of that area. The results adding to this information are somewhat lower but very similar.

Similar exercise I made for Panel 2, I include the Retiro and the three suburban monitors, results are similar than in the main results but lower due to the location of suburban stations. As those stations are outside of Madrid and close to parks.

5.4 Non-linearity in air pollution and heterogeneity analysis

It is natural to think that not all levels of pollution have the same effects. Indeed, as higher are the air pollution levels, higher are the increases in housing prices (Gruhl et al., 2022, Amani et al., 2022 and Chay & Greenstone, 2005) or the effects on work absence (Holub, Hospido & Wagner, 2021). I explore the non-linearities by running a quantile regression to explore the difference in housing prices by dividing the sample into two groups. One group includes values of air pollution lower than 40 $\mu g/m^3$, corresponding to the bottom 60 percent of the sample, and the top 40 percent has values of air pollution higher than 40 $\mu g/m^3$. Results in Table 4 show that for high pollution levels (columns 2, 4 and 6), the effects on housing prices are higher and significant, while for low levels of air pollution there are no significant effects on prices in line with no-results in Le Boennec & Salladarré, (2017) for a city with low levels of pollution like Nantes. Table 4 also displays results for longer time horizons, showing that the non-linear effect remains stable.

Table 4: Non linear effects of pollution on house prices

	(1)	(2)	(3)	(4)	(5)	(6)
	Last 3 months		Last 6 months		Last 9 months	
	Bottom 60%	Top 40%	Bottom 60%	Top 40%	Bottom 60%	Top 40%
Log NO2	-0.019	-0.125*	-0.036	-0.593***	-0.083*	-0.674***
	(0.605)	(0.080)	(0.141)	(0.000)	(0.098)	(0.000)
Observations	63,135	42,185	61,110	40,942	59,326	39,865
R-squared	0.609	0.607	0.598	0.624	0.595	0.615
House characteristics	YES	YES	YES	YES	YES	YES
Neighborhood	YES	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES	YES
Union Airbnb* year	YES	YES	YES	YES	YES	YES

The dependent variable is the logarithm of housing price per square meter. The coefficient of interest show how housing prices change when considering air pollution over the last 3 months (column 1 and 2), the last 6 months (column 3 and 4) and the last 9 months (column 5 and 6). Columns (1), (3) and (5) show the bottom 60% of air pollution, whereas columns (2), (4) and (6). All columns include house characteristics and neighborhood dummies, year and neighborhood fixed effects and group of districts times year fixed effects. Robust standard errors in parentheses are clustered by year. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Next, I check how the effect of air pollution on housing prices might be different across different types of neighborhoods. Air pollution affects more to educated people (Mikula and Pytlikova, 2021) and high income people (Cai et al., 2024; Mikula & Pitlikova, 2021). Also, migration literature has shown that people take into account air quality when they decide to migrate (Mikula & Pitlikova, 2021; Chen et al., 2022). I used the information collected in "Panel de hogares" in the webpage of the city council of Madrid to know the income, education and number of born-foreigners of each neighborhood.

Panel 1 shows the results for income, I divided the sample in two. One sample has the neighborhoods with lower income (bottom 75 percent) and the other sample the higher income (top 25 percent). In columns (1) and (2) I observe that housing prices are affected by air pollution in both low and high income neighborhoods but the coefficient is higher for high-income neighborhood, meaning that higher income individuals are willing to pay more for houses in less polluted areas. For the last six months and the last nine months, similar patterns arrise. I find similar results when differentiating neighborhodds by education. To this end I compute the share of highly educated as the percentage of individuals with a tertiary degree (see Panel 2, Table 5).

Finally regading foreigners, I differenciate between those whose birth country is outside the EU and OECD countries (Panel 3, Table 5) and those whose birth country is within the EU and OECD countries (Panel 4, Table 5). The results are similar for both groups, immigrants no matter their country of birth seem to take air pollution levels into account more than natives when buying a house. In particular, I find a stronger effect for neighborhoods with a larger share of immigrants. Also, the effects of air pollution on housing prices increases as does the time horizon. But, there is a difference between Panel 3 and 4, immigrants from EU and OECD pay more in general for clean air than Panel 3 which make sense as these immigrants are wealthier. These results are inline with the findings in the migration literature that suggest that air quality is responsible for migrantion.

Table 5: Heterogeneous effects of air pollution on house prices

	(1)	(2)	(3)	(4)	(5)	(6)
	Last 3 months		Last 6 months		Last 9 months	
	Bottom 75%	Top 25%	Bottom 75%	Top 25%	Bottom 75%	Top 25%
1. Income: Log NO2	-0.055***	-0.066***	-0.096***	-0.111***	-0.132***	-0.189***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)	(0.000)
Observations	85,306	20,014	82,712	19,340	80,385	18,806
R-squared	0.593	0.325	0.592	0.327	0.592	0.323
2. Education: Log NO2	-0.054***	-0.065***	-0.093***	-0.112***	-0.117**	-0.232***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.011)	(0.000)
Observations	83,103	22,217	80,549	21,503	78,249	20,942
R-squared	0.573	0.291	0.573	0.293	0.573	0.289
3. Foreign except EU and OECD: Log NC	-0.054***	-0.061***	-0.088***	-0.116***	-0.115**	-0.188***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.031)	(0.000)
Observations	72,528	32,792	70,153	31,899	68,109	31,082
R-squared	0.569	0.582	0.57	0.58	0.569	0.578
4. Foreign only EU and OECD: Log NO2	-0.055***	-0.072***	-0.092***	-0.132***	-0.112**	-0.266***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.017)	(0.000)
Observations	79,006	26,314	76,609	25,443	74,477	24,714
R-squared	0.577	0.548	0.577	0.546	0.576	0.543

The dependent variable is the logarithm of housing price per square meter. Each panel show a different quantile regression. For the first panel, the sample is divided by neighborhoods income: columns (1), (3) and (5) represents neighborhoods with incomes in the bottom 75 percent of the sample, while columns (2), (4) and (6) represent neighborhoods in the top 25. Additionally,the sample is divided by the percentage of residents with tertiary education. Columns (1), (3), and (5) represent neighborhoods with high education in the bottom 75 percent of the sample, while columns (2), (4), and (6) represent neighborhoods with high education in the top 25 percent. In panel 3 and 4, the percentage of foreigners is calculated by dividing between those outside from EU and OECD (panel 3) and those from within the EU and OECD (panel 4). Within each group, neighborhoods are further differentiated by the percentage of foreigners: the top 25 percent and the bottom 75 percent. The variables for income, education, and the percentage of foreigners refer to 2017 data from the neighborhood panel available on the Madrid city website. All columns include house characteristics and neighborhood dummies, year and neighborhood fixed effects and group of districts times year fixed effects. Robust standard errors in parentheses are clustered by year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6. Conclusions

This paper examines the causal effects of air pollution on housing prices, based on detailed transaction data, combined with precise air quality information for Madrid city. I use information on thermal inversion, a meteorological phenomenon, as an instrumental variable to address endogeneity concerns. My analysis shows that housing prices decrease by 0.058% when air pollution increases by 1%. The difference in air pollution between the most and the least polluted neighborhood in my sample is 22.37% and considering the average house (75 m2 and 2,723 €/m2) this implies a difference in housing prices of 2,650 euros. As buying is a time-consuming process, I compute mean air pollution over the three, six and nine months prior to each transaction, finding stronger effects for longer time horizons; the above-mentioned price differences increase to 6,396 euros when considering a nine months' horizon.

Quantile regressions indicate strong non-linearity. When air pollution is below a certain limit, there is no effect on housing prices while it is negative and significant for higher levels. In line with previous studies (Cai et al., 2024; Mikula and Pytlikova, 2021), my results show a stronger effect in neighborhoods with more high-educated and high-income individuals. Moreover, I also find stronger effects for neighborhoods with a larger share of immigrants which is aligned with findings from migration literature showing that air pollution is responsible for migration.

My study has several limitations that must be considered when interpreting the results. First, my transaction data does not provide information on either the current conditions of the house or the floor of the apartment which are important variables for determining housing prices. Second, I use interpolation techniques to estimate pollution levels and thermal inversions for each neighborhood and hence these values may not precisely reflect the actual conditions at the dwelling. An interesting road for future research could be to explore the effect of air pollution on housing prices in other Spanish cities. Additionally, incorporating subjective assessments of air quality at the neighborhood level (Montero & Fernandez-Avilés, 2013) could help to improve the model.

7. References

Amini, A., K. Nafari & R. Singh (2022) "Effect of air pollution on house prices: Evidence from sanctions on Iran" *Regional Science and Urban Economics* 93.

Anselin, L., & J. Le Gallo (2006) "Interpolation of Air Quality Measures in Hedonic House Price Models: Spatial Aspects". *Spatial Economic Analysis*, 1(1), 31–52.

Cai, Y. & M. J. Smit (2024) "Economic effects of air quality on housing prices: evidence from Beijing, China" *Journal of Housing and the Built Environment*

Chay, K. Y. & M. Greenstone (2003). The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession. *Quarterly Journal of Economics*, 118(3), 1121–1167.

Chay, K. Y. & M. Greenstone (2005) "Does Air Quality Matter? Evidence from the Housing Market", *Journal of Political Economy*, 113, 376–424,

Chen S., P. Oliva & P. Zhang (2022) "The effect of air pollution on migration: Evidence from China" *Journal of Development Economics*, Volume 156, 102833.

Cheng Keat Tang (2021) "The cost of traffic: evidence from the London congestion charge" *Journal of Urban economics,* volume 121

Cohen, A.J. & C. A. Pope (1995) "Lung cancer and air pollution" *Environ Health Perspect*,108 suppl 8:219-24.

Deryugina, T., G. Heutel, N. H. Miller, D. Molitor & J. Reif (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, 109(12), 4178–4219.

Dons, E., M. Laeremans, E. Anaya-Boig, I. Avila-Palencia, C. Brand, A. de Nazelle, M. Gaupp-Berghausen, T. Götschi, M. Nieuwenhuijsen, J. P. Orjuela, E. Raser, A. Standaert & L. Int Panis on behalf of the PASTA Consortium (2018) "Concern over health effects of air pollution is associated to NO_2 in seven European cities." *Air Qual Atmos Health* 11, 591–599.

European Commission (2019) "Attitudes of Europeans towards Air Quality" Special Eurobarometer 497.

Forsberg, B., N. Stjernberg & S. Wall (1996) "People can detect poor air quality well below guideline concentrations: a prevalence study of annoyance reactions and air pollution from traffic." *Occupational & environmental medicine* 54(1): 44–48.

Fotocasa research (2017) "Experiencia de compra y venta de vivienda en el último año".

Fredrik, N., G. Per, J. Lars, B. Tom, B. Niklas, J. Robert & P. Göran (2000) "Urban Air Pollution and Lung Cancer in Stockholm". *Epidemiology* 11(5): 487-495.

Gil, J., & J. Sequera (2022). "The professionalization of Airbnb in Madrid: far from a collaborative economy". *Current Issues in Tourism*, 25(20), 3343–3362.

Gruhl, H., N. Volkhausen, N. Pestel & N. aus dem Moore (2022) "Air pollution and the housing market: Evidence from Germany's Low Emission Zones", *Ruhr Economic Papers*, No. 977.

Guarnieri, M. & J. R. Balmes (2014) "Outdoor air pollution and asthma" *The lancet*, volume 383, issue 9928

Hanna, R. & P. Oliva (2015) The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City. *Journal of Public Economics*, 122, 68–79.

Holub, F., L. Hospido & U. J. Wagner (2020) "Urban air pollution and sick leaves: evidence from social security data", *Banco de España*, Documento de Trabajo 2041

Klomenko, S., M. Cirach, E. Pereira-Barboza, N. Mueller, J. Barrera-Gómez, D. Rojas-Rueda, K. de Hoogh, G. Hoek PhD & M. Nieuwenhuijsen (2021) "Premature mortality due to air pollution in European cities: a health impact assessment", *The Lancet Planetary health*, volume 5 issue 3, e121-e134

Le Boennec, R. & F. Salladarré (2017) "The impact of air pollution and noise on the real estate market. The case of the 2013 European Green Capital: Nantes, France," *Ecological Economics*, vol. 138(C), pp. 82-89.

Mikula, S. & M. Pytlikova (2021). " Air Pollution and Migration: Exploiting a Natural Experiment from the Czech Republic" *IZA Discussion Papers* 14863, Institute of Labor Economics (IZA).

Montero, JM. & G. Fernández-Avilés (2013) "Measuring the impact of pollution on property prices in Madrid: objective versus subjective pollution indicators in spatial models. " J Geogr Syst 15, 169–191.

Moral- Carcedo, J. (2023) "Urban growth and stratification: the role of locational externalities", mimeo.

Appendix A

Figure A1. Districts and neighborhood of Madrid city

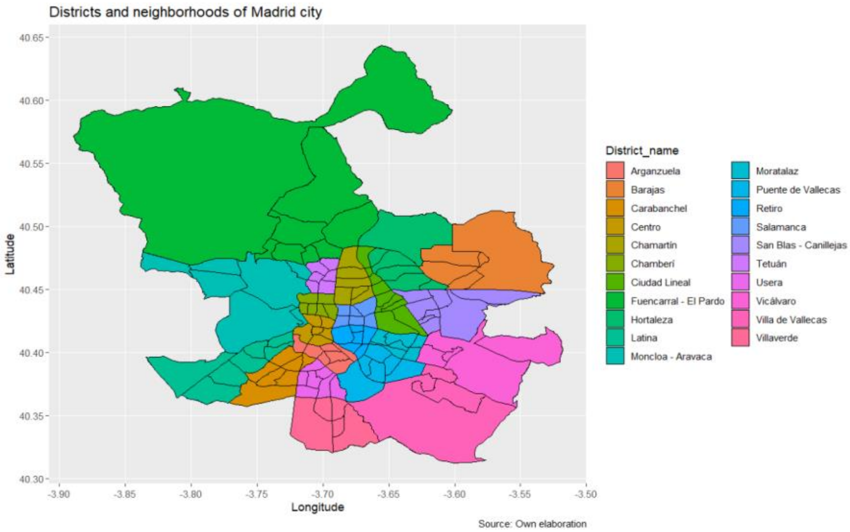

Figure A2. Map of pollution stations of Madrid neighborhoods.

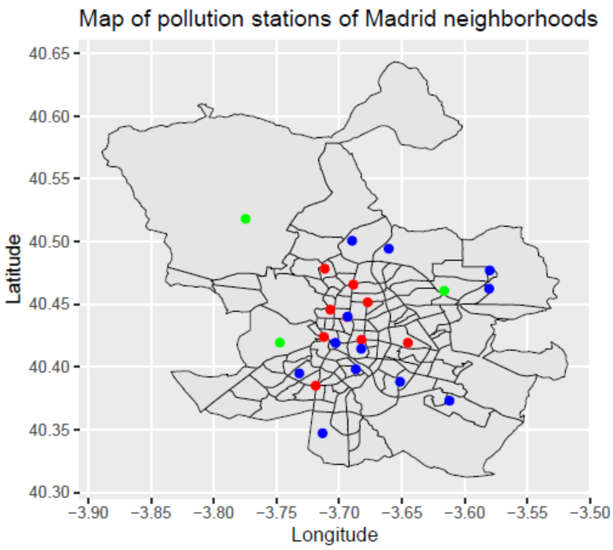

Figure A3. Thermal inversion.

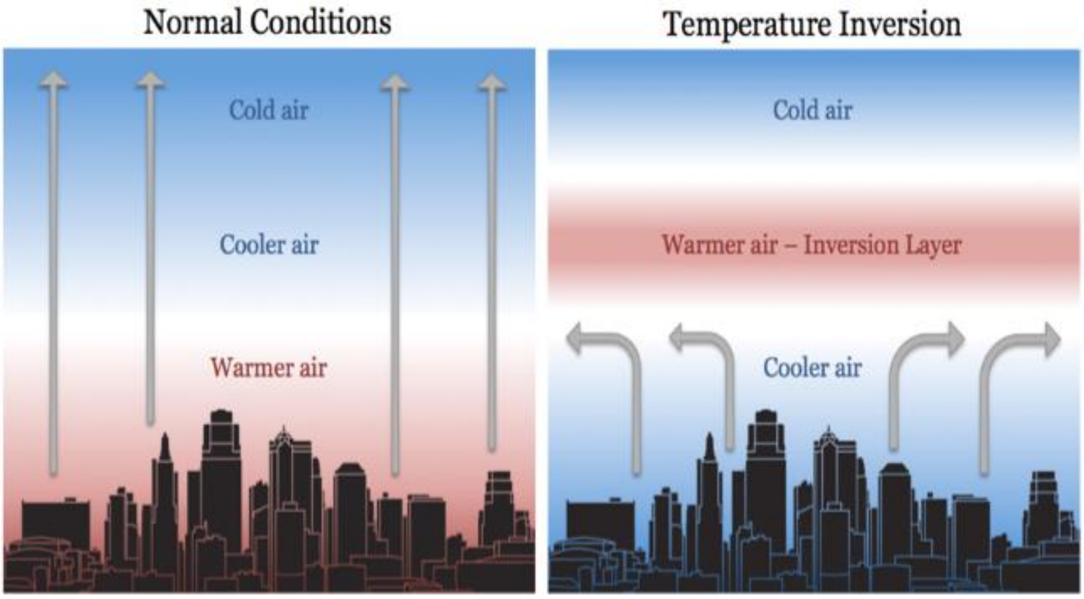

Figure A4. Variation of house prices between 2016 and 2017.

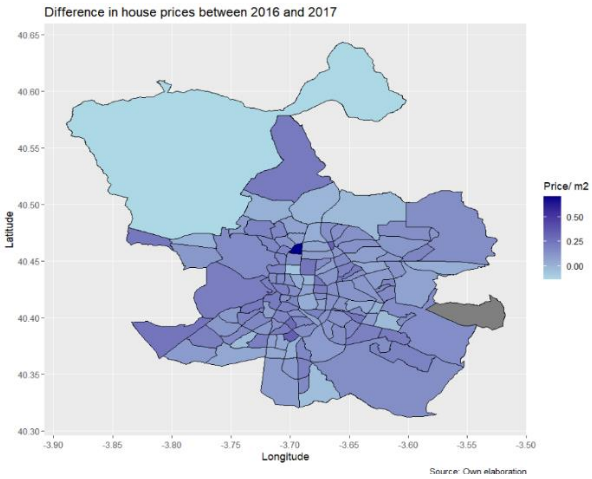

Figure A5. Variation of air pollution between 2016 and 2017.

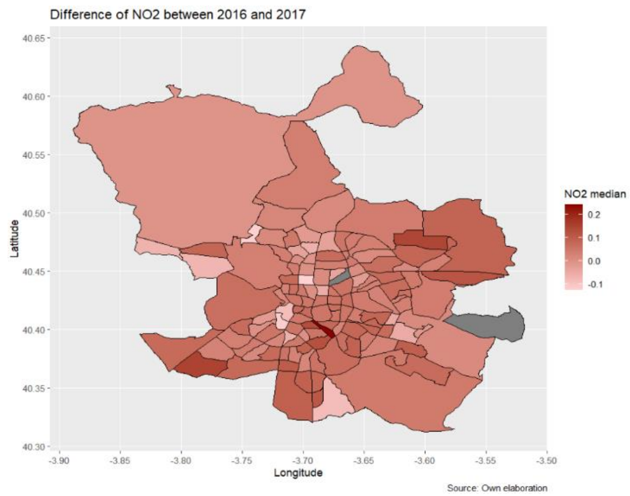

Figure A6. Variation of thermal inversion between 2016 and 2017.

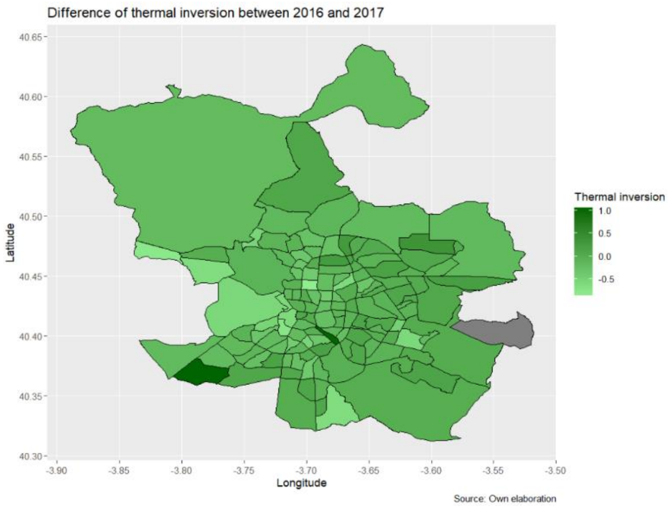

Table A1: Airbnb concentration by districts

Districts	Airbnb totals	Entire home	Nº households	% total Airbnb	% entire home
Centro	9,334	6,881	69,504	13.43	9.9
Salamanca	1,419	1,113	63,001	2.25	1.77
Chamberí	1,257	802	61,936	2.03	1.29
Arganzuela	1,078	658	65,479	1.65	1
Retiro	747	470	48,880	1.53	0.96
Tetuán	984	640	67,073	1.47	0.95
Chamartín	603	406	58,413	1.03	0.7
Moncloa - Arav	604	306	46,361	1.3	0.66
Carabanchel	732	338	96,871	0.76	0.35
Usera	353	177	50,285	0.7	0.35
Ciudad Lineal	642	270	87,225	0.74	0.31
Barajas	158	54	18,756	0.84	0.29
San Blas - Canil	472	176	60,551	0.78	0.29
Hortaleza	446	211	72,104	0.62	0.29
Puente de Valle	564	245	89,291	0.63	0.27
Latina	550	198	95,756	0.57	0.21

The data on AIRBNB is from their webpage on september of 2022 and the data of number of household is from the city council of Madrid webpage.

Table A2: Estimation results for OLS regression including controls

		(1)	(2)	(3)	(4)	(5)
	Log NO2	0.093**	0.042***	-0.060***	-0.045***	-0.047***
Year of construction	Less than 1960		0.092***	-0.319***	-0.313***	-0.314***
	Between 1960-1975		-0.241***	-0.284***	-0.279***	-0.279***
	Between 1975 -2000		-0.204***	-0.230***	-0.226***	-0.227***
	old		-0.160***	-0.082***	-0.08***	-0.085***
Property Type	Flat without building		-0.157***	-0.085***	-0.087***	-0.088***
	Semi-detached		-0.131***	0.102**	0.090*	0.089*
	Detached		-0.089***	0.091*	0.094*	0.093*
Property size	Less than 55m2		-0.141***	0.161***	0.163***	0.163***
	Between 55- 75m2		-0.290***	0.067***	0.067***	0.067***
	Between 75-100 m2		-0.243***	0.019	0.021	0.021
Activities CNAE	Mean of human health			0.006*	-0.003*	-0.003**
	Mean education			0.000	0.002**	0.002
	Mean wholesale and retail			-0.001***	0.000	0.000
	Mean accomodation			0.006***	0.000	0.000
	Mean other services			0.005**	0.002***	0.002***
	Mean entertainment and recreation			0.014***	0.001	0.001
	Total population			0.000	-0.000***	-0.00***
Observations		105,629	105,629	105,320	105,320	105,320
R-squared		0.002	0.138	0.574	0.606	0.607
House characteristics		NO	YES	YES	YES	YES
Neighborhood		NO	NO	YES	YES	YES
Year		NO	NO	NO	YES	YES
Union Airbnb* year		NO	NO	NO	NO	YES

The dependent variable is the logarithm of housing price per square meter. In column (1) no controls are included, in column (2) house characteristics dummy controls are included, in column (3) neighborhood controls and fixed effects and in column (4) fixed efects are included. Finally, column (5) has all controls and it is the baseline model . Robust standard errors in parentheses are clustered by year in column (4) and (5). $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$.

Table A3: Estimation results for IV regression including controls

		(1)	(2)	(3)	(4)	(5)
	Log NO2	0.055***	0.053***	-0.012	-0.058***	-0.058***
Year of construction	Less than 1960		0.082***	-0.320***	-0.312***	-0.313***
	Between 1960-1975		-0.251***	-0.285***	-0.278***	-0.278***
	Between 1975 -2000		-0.214***	-0.231***	-0.226***	-0.226***
	old		0.295***	0.034***	0.028***	0.027***
Property Type	Flat without building		-0.157***	-0.085***	-0.087***	-0.088***
	Semi-detached		-0.128***	0.104**	0.090***	0.089***
	Detached		-0.089***	0.088*	0.091***	0.091***
Property size	Less than 55m2		-0.143***	0.160***	0.161***	0.161***
	Between 55- 75m2		-0.291***	0.067***	0.067***	0.067***
	Between 75-100 m2		-0.243***	0.018	0.020***	0.020***
Activities CNAE	Mean of human health			0.006*	-0.003***	-0.003***
	Mean education			-0.002	0.002***	0.001**
	Mean wholesale and retail			-0.001***	0.000	0.000
	Mean accomodation			0.005***	0.000	0.000
	Mean other services			0.006**	0.002***	0.002***
	Mean entertainment and recreation			0.013**	0.000	0.001
	Total population			0.000	-0.000***	-0.000***
Observations		105,629	105,629	105,320	105,320	105,320
R-squared		0.001	0.137	0.574	0.606	0.607
House characteristics		NO	YES	YES	YES	YES
Neighborhood		NO	NO	YES	YES	YES
Year		NO	NO	NO	YES	YES
Union Airbnb* year		NO	NO	NO	NO	YES

The dependent variable is the logarithm of housing price per square meter. In column (1) no controls are included, in column (2) house characteristics dummy controls are included, in column (3) neighborhood controls and fixed effects and in column (4) fixed efects are included. Finally, column (5) has all controls and it is the baseline model . Robust standard errors in parentheses are clustered by year in column (4) and (5). $*\ p<0.10$, $**\ p<0.05$, $***\ p<0.01$.

Table A4: Regression results after eliminating districts

	(1) Last 3 months	(2) Last 6 months	(3) Last 9 months
Panel 1: Without district 1			
Log NO2	-0.056***	-0.098***	-0.137***
	(0.000)	(0.000)	(0.001)
Observations	97,602	94,584	91,929
R-squared	0.604	0.604	0.603
Panel 2: Without district 18 and 19			
Log NO2	-0.058***	-0.100***	-0.144***
	(0.000)	(0.000)	(0.000)
Observations	101,821	98,768	96,107
R-squared	0.608	0.607	0.606
House characteristics	YES	YES	YES
Neighborhood	YES	YES	YES
Year	YES	YES	YES
Union Airbnb* year	YES	YES	YES

The dependent variable is the logarithm of housing price per square meter. For panel 1, district 1 is eliminated from the sample and in panel 2 districts 18 and 19 are eliminated from the sample. All columns include house characteristics and neighborhood dummies, year and neighborhood fixed effects and group of districts times year fixed effects. Robust standard errors in parentheses are clustered by year. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A5: Regression results after controlling by traffic

	(1) Last 3 months	(2) Last 6 months	(3) Last 9 months
Log NO2	-0.058***	-0.096***	-0.135***
	(0.000)	(0.000)	(0.001)
Mean occupation	0.000	-0.001	-0.001
	(0.603)	(0.212)	(0.392)
Observations	101,718	98,541	95,768
R-squared	0.597	0.597	0.596
House characteristics	YES	YES	YES
Neighborhood	YES	YES	YES
Year	YES	YES	YES
Union Airbnb* year	YES	YES	YES

The dependent variable is the logarithm of housing price per square meter. The coefficient of interest show how housing prices change when considering air pollution over the last 3 months (column 1), the last 6 months (column2) and the last 9 months (column 3). Traffic control is included in this regression, the data on traffic is provided by the city council of Madrid. All columns include house characteristics and neighborhood dummies, year and neighborhood fixed effects and group of districts times year fixed effects. Robust standard errors in parentheses are clustered by year. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A6: Regression results using other measures of air pollution

	(1) Last 3 months	(2) Last 6 months	(3) Last 9 months
Panel 1: Using mean of air pollution			
Log NO2	-0.065***	-0.112***	-0.164***
	(0.000)	(0.000)	(0.000)
Observations	105,320	102,052	99,191
R-squared	0.607	0.607	0.606
Panel 2: Using other technique of interpolation			
Log NO2	-0.055***	-0.095***	-0.135***
	(0.000)	(0.000)	(0.003)
Observations	101,718	98,541	95,768
R-squared	0.607	0.607	0.606
House characteristics	YES	YES	YES
Neighborhood	YES	YES	YES
Year	YES	YES	YES
Union Airbnb* year	YES	YES	YES

The dependent variable is the logarithm of housing price per square meter. For panel 1, instead of using the monthly median, the monthly mean of air pollution is used. In panel 2, instead of using Kriging interpolation for polution, the inverse distance- weighted average of the three closest monitors for each neighborhood is used. All columns include house characteristics and neighborhood dummies, year and neighborhood fixed effects and group of districts times year fixed effects. Robust standard errors in parentheses are clustered by year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Regression results after including Retiro and suburban monitors

	(1) Last 3 months	(2) Last 6 months	(3) Last 9 months
Panel 1: Including Retiro monitor			
Log NO2	-0.057***	-0.098***	-0.142***
	(0.000)	(0.000)	(0.000)
Observations	105,320	102,052	99,191
R-squared	0.607	0.607	0.606
Panel 2: Including Retiro and suburban monitors			
Log NO2	-0.056***	-0.097***	-0.141***
	(0.000)	(0.000)	(0.000)
Observations	105,320	102,052	99,191
R-squared	0.607	0.607	0.606
House characteristics	YES	YES	YES
Neighborhood	YES	YES	YES
Year	YES	YES	YES
Union Airbnb* year	YES	YES	YES

The dependent variable is the logarithm of housing price per square meter. For panel 1, Retiro monitor is included and for panel 2 Retiro and suburban monitorsa are included. All columns include house characteristics and neighborhood dummies, year and neighborhood fixed effects and group of districts times year fixed effects. Robust standard errors in parentheses are clustered by year. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix B

The Spanish legislation on air quality is closely related to the European regulation, current legislations about this topic are: Law 34/2007, that updates the legal basis for developments related to the assessment and management of air quality in Spain and has the aim of achieving optimum levels of air quality. Also, a Royal Decree 102/2011, this regulation transposes the content of two European Directive. This royal decree was amended by other two Royal Decree, and provides the approval of a National Air Quality Index to inform citizens, in a clear and homogeneous manner throughout the national territory, about the quality of the air they breathe at any given moment.

The National index defines 6 categories of air quality: good, reasonably good, fair, unfavorable, very unfavorable, and extremely unfavorable. Each station is assigned the worst category in terms of air quality for any of the pollutants considered for estimation. The pollutants considered in the index are: Suspended Particulate Matter smaller than 10 micrometers (PM_{10}), Suspended Particulate Matter smaller than 2.5 micrometers ($PM_{2.5}$), Tropospheric Ozone (O_3), Nitrogen Dioxide (NO_2) and Sulfur Dioxide (SO_2). In this research, I use the level of NO_2 as an indicator for air pollution in Madrid city, as it is one of the main pollutants in the national index, but also United States Environmental Protection Agency and World Health Organization has NO2 in their pollution index. Moreover, there are 24 air pollution monitors that measure this pollutant, whereas for PM_{10} and Ozone there is 13 monitors, $PM_{2.5}$ has 8 monitors and SO_2 has 4 monitors. Finally, NO_2 is a gas that is emitted in the combustion processes carried out by motor vehicles and in high temperature industrial and electricity generation, it is easily recognized by its reddish-brown color and its characteristic pungent odor. The air quality directive has set two limit values for nitrogen dioxide (NO_2) for the protection of human health: the NO_2 hourly mean value may not exceed 200 micrograms per cubic meter (µg/m^3) more than 18 times in a year and the NO_2 annual mean value may not exceed 40 µg/m^3.