Utrecht University

The Impact of Airbnb on the Short-Term Rental Market: evidence for Italy

Enrico Bergamini
6313035

Master's Thesis
MSc in Economic Policy
Academic Year 2017/2018

Table of Contents

Abstract

This research investigates the effect that Airbnb's presence has on the short-term rental market in seven Italian cities: Bologna, Florence, Milan, Palermo, Rome, Torino and Venice. Performing a visual analysis confirms that Airbnb is highly concentrated in the historic centers, especially in tourism-intense cities like Venice and Florence. In order to estimate the impact on rental rates, I make use of two different empirical strategies. Firstly, I regress Airbnb's presence on rents using a dataset built with public data. Secondly, I make use of hedonic regressions on highly granular web-scraped datasets. In both cases, I find evidence that Airbnb's presence is positively and significantly associated with higher short-term rental rates.

Keywords: airbnb, real estate economics, spatial econometrics, QGIS, data science, gentrification, urban political economy, hedonic regressions

1. Introduction

Airbnb, Uber, Lyft and GetAround are some of the pioneers of the sharing-economy: a form of economic activity that is based on disintermediation and peer-to-peer platforms. Under this framework, economic agents can be at the same time consumers and producers, as famously captured by Jeremy Rifkin with the term "*prosumers*".[1]

Airbnb was founded in 2008: it is a platform on which hosts could make their private housing spaces available for rent for a short-term period. The idea started in San Francisco in the fall of 2007, during a design-festival that overcrowded the city, and sold-out hotel rooms. Brian Chesky and Joe Gebbia, the two co-founders of Airbnb (at that time sharing a house) decided to buy three air mattresses and rent them for 80$ per night to help cover their own rent expenses.[2]

Airbnb's growth has been explosive since its foundation: figures for 2017 show 2.6 billion revenues, and almost 100 millions in profits.[3]
Airbnb is now present in 81000 cities across 191 countries, and is used by more than 300 million guests. For 2017 it records nearly 5 million listings registered on the platform,

from couches to entire apartments and buildings, including also 3000 castles and 1400 treehouses.[4]

The figures are impressive, and the company defines itself as a *"global travel community that offers magical end-to-end trips, including where you stay, what you do and the people you meet"*, being able to *"economically empower millions of people around the world to unlock and monetize their spaces, passions and talents to become hospitality entrepreneurs"*.[5]

However, taking into consideration the size and worldwide penetration that Airbnb has reached we can pose ourselves the question: what is the bigger impact of these numbers?

The user-empowering side of the story is certainly part of the truth, but it is not the only one: ten years after the explosion of platform-economy, it has become clear that this is a tale of two narratives (Pasquale, 2016).[6]

While the arguments for efficiency, disintermediation, cultural exchange and sustainability are proposed by the sharing-economy advocates, a body of critiques has been growing. While its explosion certainly exposed an underlying demand, it is also provoking the touristization of city-centers, a phenomenon that is arising discussions in many cities.[7] Some argue that city centers are rapidly being transformed into "theme parks for tourists",[8] and local long-term residents are more often suffering increased nuisance and bothering.[9]
Moreover, as I will discuss in further detail in the next Section, many hosts are *de facto* landlords listing several properties for full-time rental to tourists, profiting from an unregulated industry.

Thus, what are the effects of home-sharing on the urban economies? What is the aggregate effect on welfare? How should policymakers react? This broader set of social concerns is what Pasquale (2016)[10] calls the *progressive narrative*, which opposes the techno-optimist and neoliberal view, typically more concerned with competition, disruptive innovation and market freedom.

Among economists and social scientists, there is an increasing attention on the broad impact of platforms like Airbnb, in terms of gentrification and touristization. While the peer-to-peer promise is real and appealing for many on the platform, there is concrete risk to increase inequality and short-sell our cities, if left unregulated. In the optic of protecting social welfare, the economic reality behind the use of platforms like Airbnb must be assessed.

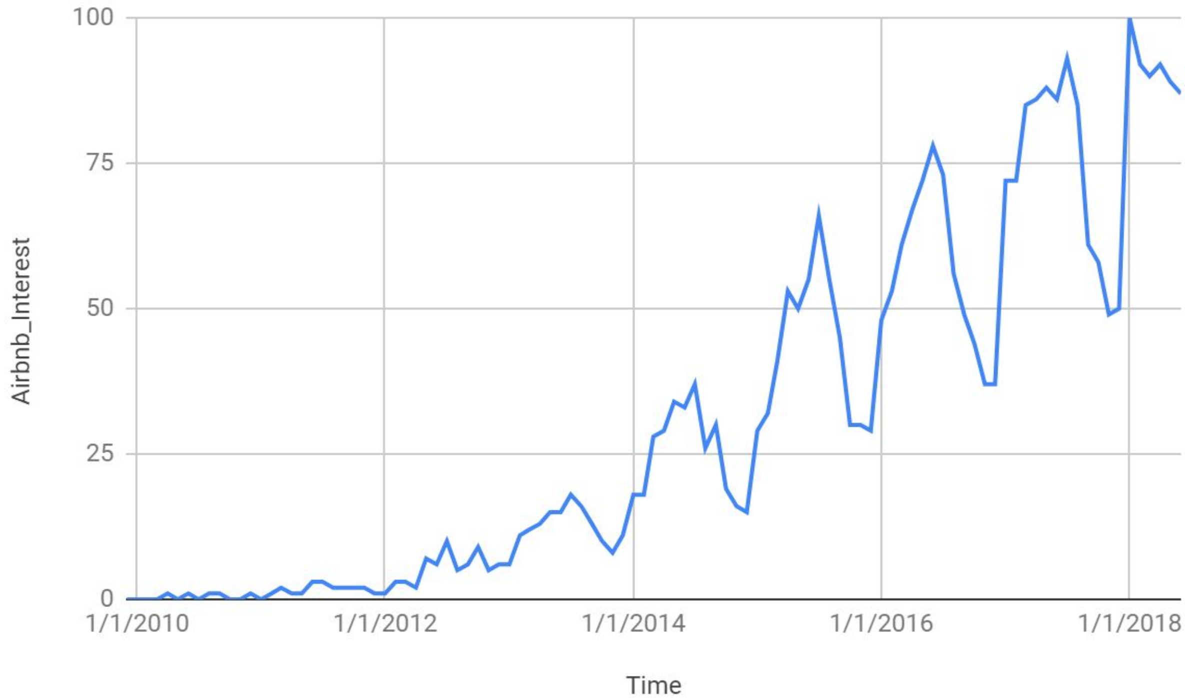
A small but increasing body of knowledge is focusing on the influence that Airbnb has on housing affordability, due to its effects on the short-term rental rates (STR). In fact, the preference for Airbnb rental over traditional STR decreases the supply of housing available, therefore rising the prices in the area.

This and many other concerns will be discussed in the next Section in further detail.
My research aims at shedding more light on this debate by an empirical assessment of seven Italian cities: Bologna, Florence, Milan, Palermo, Rome, Torino and Venice.

Figure 1: Cities considered

I will focus on the Italian real estate market as it presents features of interest in terms of my research question. The following graph shows the worldwide interest in the search term "Airbnb Italy" according to Google Trends data. The peaks in the graph are in the summer months.

Figure 2: Worldwide interest in the term "Airbnb Italy" for the period 2009-2018



Italy, in fact, is among the top 10 inbound countries for Airbnb guests, as stated by the company itself.[11]

In addition, according to the Tourism Satellite Account computed by ISTAT (the Italian National Institute for Statistics) in 2015, the value added produced by the Italian tourism-related industry amounted to 87.8 millions (6% of the overall value-added in the economy).[12]

The combination of high Airbnb density and a strong tourism-related industry make Italy a very interesting market to assess the impact of home-sharing on urban economies.

Although the single-impact of Airbnb might be relatively low, I would also like to stress the importance of understanding housing-price dynamics for the Italian society. Seeing the general trends in disposable income and housing prices, PWC's Real Estate Market Overview,[13] highlights a positive trend for post-2008 recovery both in investments and disposable income, with Rome and Milan driving the market.
However, inequality and inter-generational dynamics are concerning,[14] and one of the structural causes for the low rate of emancipation of the Italian youth is a combination of poor labor market conditions and high housing costs.[15]

Hence, the relevance of my research question: **what is the impact of Airbnb's presence on the short-term rental market?**

The document is organized as follows.
First, I will provide a literature review. In Section 3 I present the Airbnb dataset. Then, I will make use of two distinct empirical approaches to the question, in Sections 4 and 5. For each empirical approach, I will make use of different datasets and models, which will be described and discussed in each section. I will then discuss the respective results and limitations.
Section 6 concludes. The reader can find an Appendix at the end of this document with a comprehensive collection of tables and figures.

2. Literature Review

The impact of Airbnb has been tested in a number of cities across the world. This stream of literature is small but growing, with different angles from which to observe the issue: the effect of Airbnb on short-term rental (STR) and house prices, on the hotel industry and on local economies and neighbourhoods.

The most extensive work (in terms of markets considered) is that of Barron, Kung and Proserpio (2017).[16] Their research covers the entire United States, with a timespan that goes from 2012 to 2016.

I referred to their methodology in order to construct my first empirical strategy. Their mathematical modelling entails that there are two channels of transmission by which Airbnb affects real estate prices.

In the first channel, the home-sharing platforms make landlords change their preference towards the short-term rental market vis-à-vis long-term, by reducing their costs.
They argue that higher rental rates, made possible by switching to the STR, are then reflected on house prices via a decreased supply: every house listed on Airbnb is one less in the STR market.
For the second channel, they argue that the ability to monetize extra-capacity can increase the relative value of ownership compared to rental services, therefore increasing sales prices. They find empirical proof for both channels. In particular, the growth of Airbnb over the 2012-2016 timespan is able to explain 0.27% of the annual

rent growth. Their granularity choice and data stratification methodology will be further discussed in the next Section.

Picascia, Romano and Teobaldi (2017),[17] focus on 13 Italian cities in order to plot the relationship between tourism, Airbnb presence and inequality. Their findings are highly relevant for my research, as they seem to suggest that tourism is the main driver for Airbnb supply, legitimizing the concern for rising rates faced by local residents and possibly a gentrification mechanism driving the local population out of the centers.

They look at inequality from two different perspectives: between hosts and between city areas (spatial income inequality). In both cases revenues from home-sharing are highly concentrated: *"In most cities a handful of operators listing several properties are capable of amassing more than two thirds of the total revenues generated on the website".*
The Gini indexes calculated, in fact, are as high as 0.70 in Milan, 0.62 in Rome, and 0.60 in Venice. Indeed, inequality index for these markets is above 0.50 for all the cities, and has been rising between 2015 and 2016. This confirms, despite Airbnb's narrative, that many hosts resemble more real-estate entrepreneurs yielding high returns to capital from a number of properties and lower taxes, than peer-to-peer enthusiasts. In terms of spatial income inequality they also find that a "*disproportionally high share of Airbnb earnings flow towards the central areas of cities*".

Lee (2016) investigates the role of Airbnb on Los Angeles' affordable housing crisis.[18]
His research shows signs of an ongoing gentrification process that concerns many of Los Angeles' neighbourhoods. This process tends to exasperate racial and socio-economic inequality via a sharp increase in STR rates in a city where housing is already over-burdening households' budget. One of the striking points made in his paper is that many of Airbnb's listings are of the 'entire-place' kind, meaning that there is in fact an STR market for tourists that is overlapping with the traditional STR market for the locals.

He also suggests that the touristization of neighbourhoods is associated with a lower quality of life and calls for more equitable housing policies.

Quattrone et al. (2016)[19] study the explosion of the peer-to-peer housing platforms, and the lack of regulatory frameworks that effectively tackle this phenomenon. They propose arguments for algorithmic regulation, meant to give real-time, detailed policy responses.

Dudàs et al. (2017),[20] have analyzed the data for Budapest, in a descriptive study that visualizes different indicators of Airbnb presence with the respect to different areas of the city.

Horton (2016),[21] finds a negative impact of Airbnb on housing affordability for a number of different areas of Toronto.
Horn and Merante (2017),[22] exploit a big data approach to perform a similar research on the city of Boston, finding that a one standard deviation increase in Airbnb's presence is associated with a 0.4% increase in STR rates.

As I already mentioned in the introduction, one of the reasons for concern is that of gentrification: the process through which local communities are displaced from their neighbourhoods due to the inflow of wealthier households.
Proof of gentrification via sharing-economy is found in New York by Wachsmuth et al. (2018)[23] and Cox (2017)[24]. Opposite evidence for the same city is found by Coles et al. (2017)[25]: they argue that Airbnb's impact on housing affordability is not major, and discuss several different approaches to regulation: from night and unit caps to fees and taxes. Zervas and Proserpio (2014), estimate the impact of Airbnb on the hotel industry in Texas.[26]

Their research contributes to broaden the focus of the analysis on the effect of the sharing economy on social welfare: the hotel industry is affected negatively, especially for low-budget hotels (which are a substitute), compared with high-end hotels. The effects on social welfare are ambiguous.
Coyle & Yeung (2016), investigate the same issue in 14 European cities, and suggests that the effect on the hotel industry might not be as detrimental for the hotel industry.[27]
Evidence of racial discrimination in the sharing-economy is found by Edelman et al. (2017).[28]

Generally, although Airbnb is only an accelerating factor part of a broader trend of increasing urban inequalities,[29] if lower-income households are increasingly being driven out of their neighbourhood, the phenomenon has multiple repercussions on inequality. By contributing to gentrification, it can hinder the communities' capability for social cohesiveness and the process of creating socially and environmentally sustainable cities.[30]

3. Airbnb Dataset

Airbnb does not share publicly any of its data, and this contributes to make the phenomenon difficult to observe and quantify, from the policymakers' perspective. Researches have overcome this limitation by building web scrapers. Web scrapers are softwares that automate the collection of information over a large number of web pages, by continuously crawling over them.

Most of the researches quoted during the previous Section, make use of the datasets provided by mainly three sources: InsideAirbnb,[31] Tom Slee[32] and AirDNA.[33]
The former is a civic initiative developed by Andy Murray, who made the scraped datasets publicly available. Tom Slee is an independent researcher: the data he scraped is publicly available and as well his code.[34] AirDNA, instead, is a private firm that sells scraped Airbnb's data.

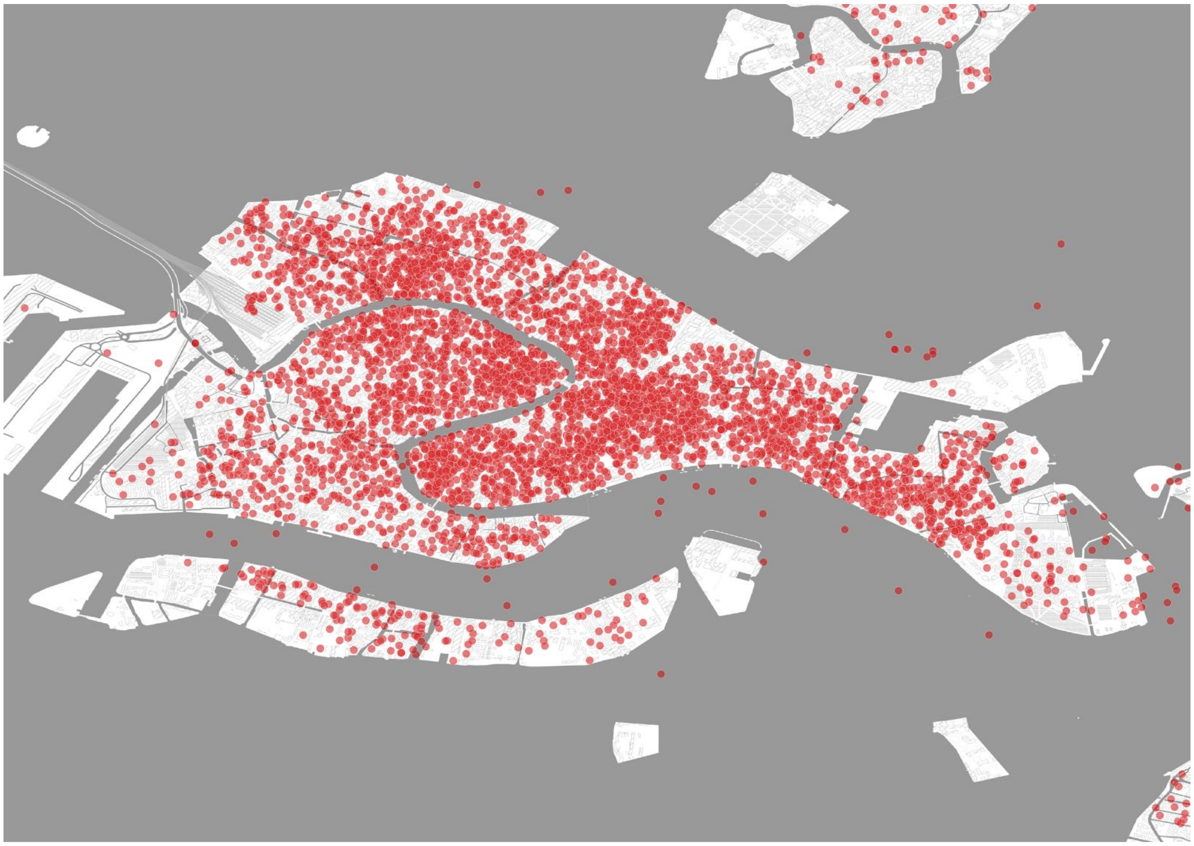
The dataset I use for my analysis is coming from a novel source.
It was provided to me by Vincenzo Patruno:[35] activist of the group OpenPuglia[36] and the association OnData.[37]

The novelty consists in the scraping technique implemented by Patruno. In fact, while the quoted sources scrape all the listings for a given city, Patruno's scraper uses a "brushing" methodology that over time collects the information about the listings that are actually active.

My analysis focuses on seven Italian markets: Bologna, Florence, Milan, Palermo, Rome, Turin and Venice and contains the active listings for the beginning of 2018.
For each listing, the dataset presents information about the type of property, the capacity (number of beds) and the number of reviews received.

Clearly, Airbnb is applying a random error factor to the listings present on the website, in order to pollute the data obtained by scraping. An example is illustrated in the picture below, in which I plot the Airbnb listings present in Venice: some of them are artificially moved into the sea. Correcting for random errors in seven cities would require a punctual crosscheck of the geolocation of each listing, a procedure that is above the scope of this research.

Figure 3: Airbnb listings in the center of Venice



4. Empirical Strategy I

4.1 Datasets and data stratification

This procedure follows that implemented by Barron et al. (2017).[38]
Their dataset is constructed at on micro-areas for the whole United States, and aims at controlling for the effects of the single zip-codes (which is their smaller geographical dimension).

I use a similar approach, by layering the data to the smallest possible region, for each city, using the geographical open-source software QGIS. The major limitation of this dataset is that, compared to Barron et al. (2017), I build a cross-sectional dataset instead of a panel. The second limitation is a lack of available public data to control for income at a suburban level. I will discuss in further detail these limitations at the end of the Section.

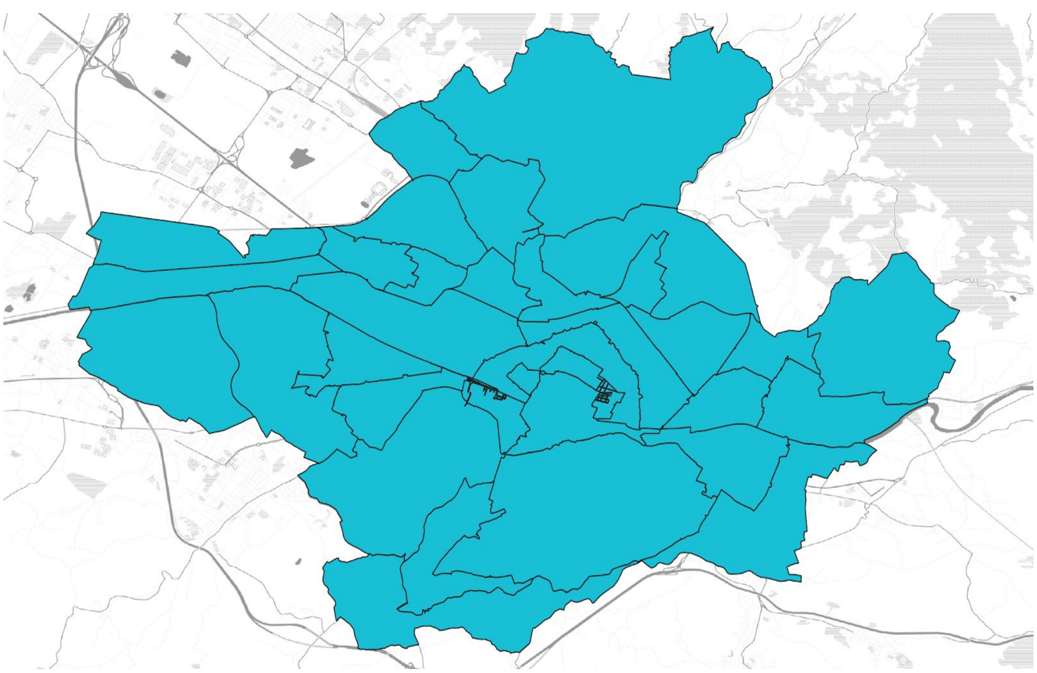
I use three data sources for all the seven cities, which have three different levels of granularity. The following table summarizes the sources:

Source	**Data**	**Granularity**
Agenzia delle Entrate	OMI **areas** and OMI **zones**: **prices** for 2017	Dozens per city (polygons)
ISTAT	CENSUS tracts: variables in stocks for **demographics** and **real estate conditions**	Hundreds or thousands per city (polygons)
Airbnb (Patruno's scraping)	Geographical presence of active hosts, reviews and capacity	Single Listings (points)

The first source is the "*Osservatorio del Mercato Immobiliare*" (OMI, onwards) by Agenzia delle Entrate, the Italian Tax Authority. OMI database is provided in open access for 2016-2017 and contains information about the average rental rates for suburban areas. There are dozens of these areas for each city.

The following map is an example of the OMI zones for Florence. Each polygon in the map contains two prices for the STR rates: minimum and maximum.

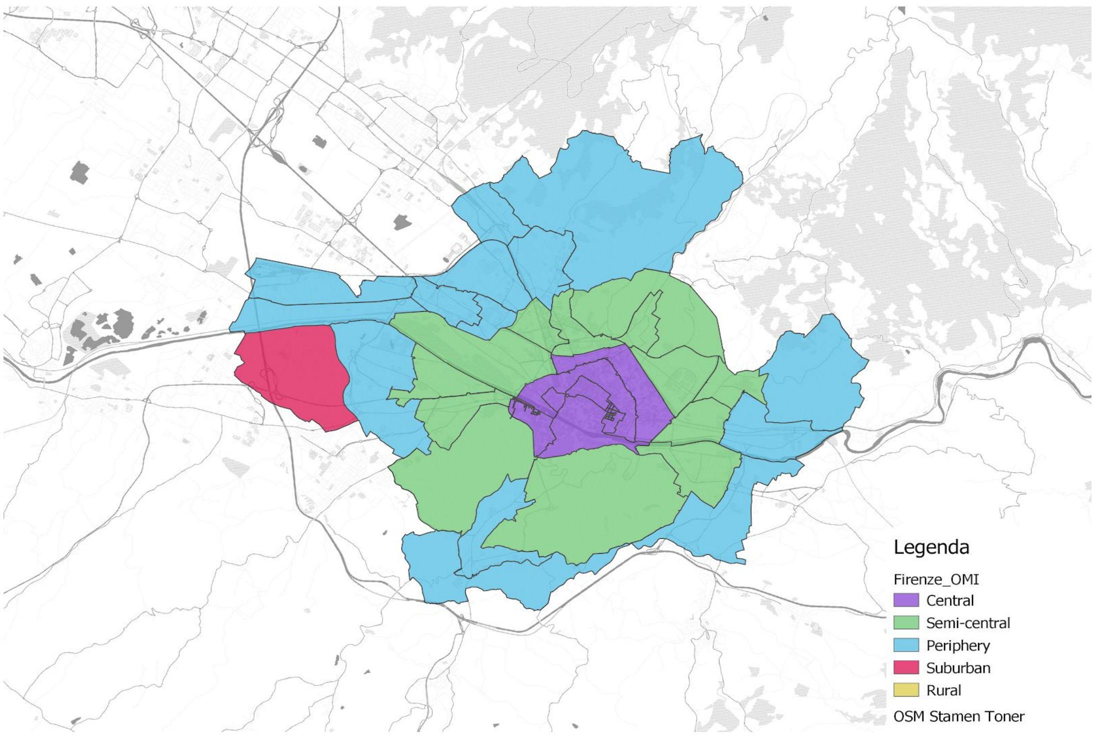
Figure 4: OMI areas in Florence



OMI zones are defined as areas in which conditions of the real estate market are homogeneous.[39]
OMI's rental rates are estimated by the institute for every semester, and I used the most recent data for the second semester of 2017.

OMI zones are also clustered into different "rings", that for each city distinguish between 'central', 'semi-central', 'periphery', 'suburban area' and 'rural area', as shown in the following map:

Figure 5: OMI rings in Florence



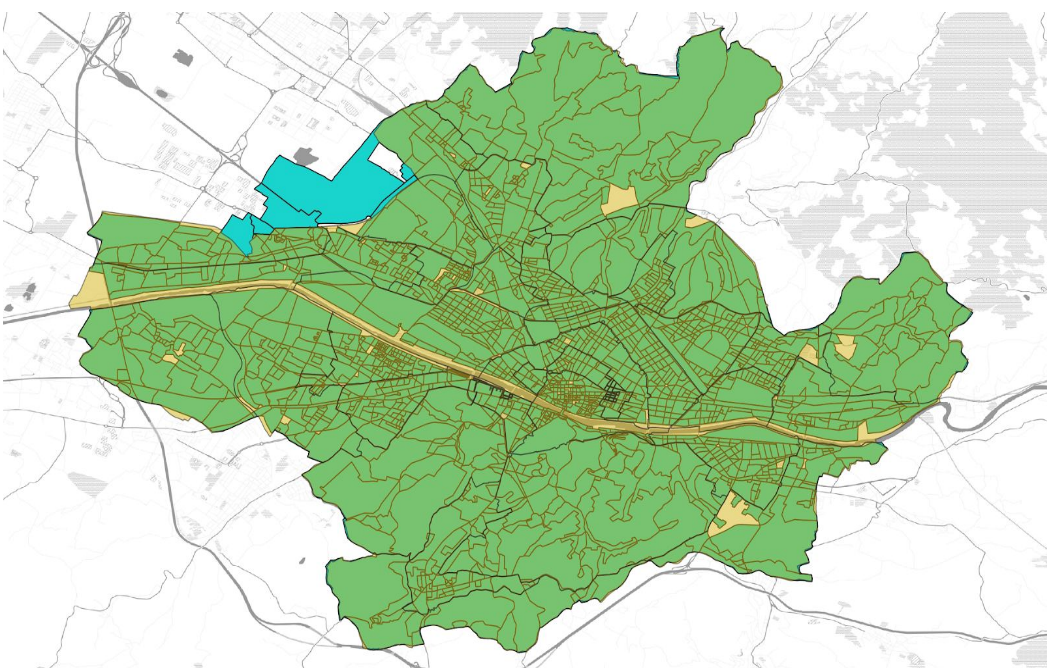
The second source is the Italian Institute for Statistics (ISTAT) that publishes microdata of the national Census of 2011 on demographics and real estate conditions, counting hundreds of polygons per city. In the map below, I show the granularity of this dataset.

Figure 6: ISTAT census tracts in Florence

I decided to stratify the data at the ISTAT level, in order to obtain a dataset in which my observations are the hundreds of ISTAT census areas.

However, the two shapefiles are not built to be precisely correspondent: many of the polygons in the ISTAT shapefiles are cut in multiple parts by the OMI zones containing average rental prices, as shown by a simple overlaying in the following map:

Figure 7: Intersection between OMI and ISTAT areas in Florence

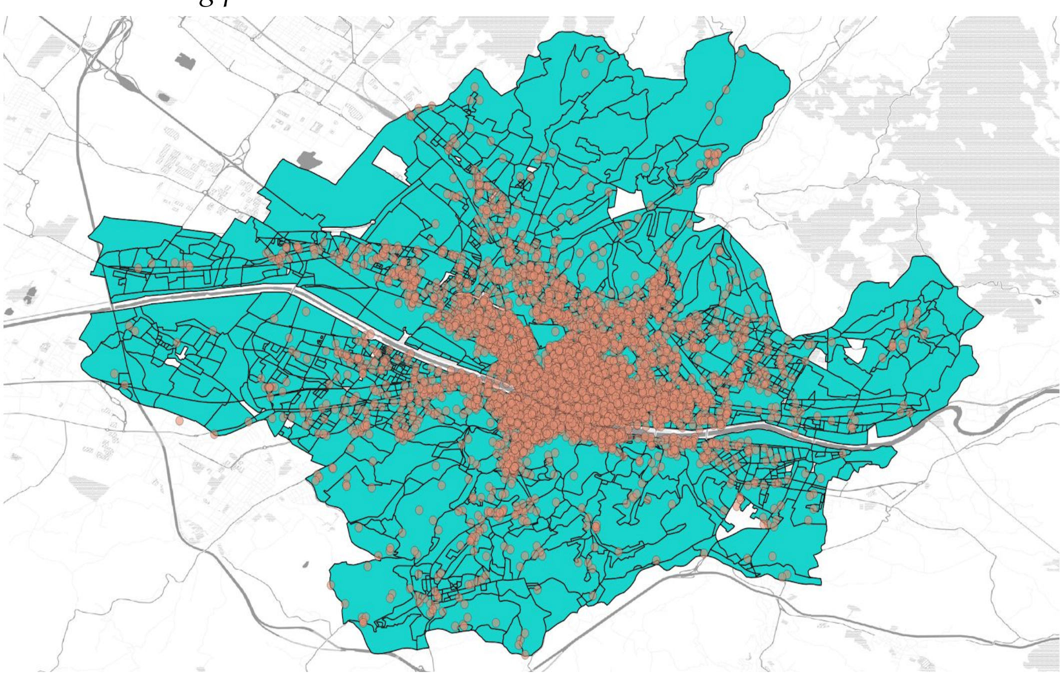


I decided to calculate the average price for each of the ISTAT census areas as a proportion of the overlapping polygons that intersect it.
To clarify, I will make an example.
Say that an ISTAT polygon is split in two, in portions of 30% and 70%, by two different OMI polygons (with two different prices). The resulting price in the ISTAT area will be an average weighted on the portions of the area covered by the OMI areas. Obviously, I am assuming that the prices are homogeneously distributed but the variation in prices is not so large to create a concern.

Therefore, in order to calculate average rental price for each ISTAT micro-area, I intersected the two layers 'cutting' them onto each other.
In this way, I obtained a geographical dataset that was composed of one or more polygons with the same ISTAT unique identifier.
After calculating the single areas of the intersected polygons, I calculated the area-weighted average price based on the unique ISTAT identifiers, I drop the duplicates, and merge the data back to the original ISTAT areas.

I obtained a layer of polygons identical in shapes to the ISTAT census tracts, but with OMI variables in addition to the ISTAT ones. At this point, I could plot the Airbnb dataset on this layer, as shown in the following map.

Figure 8: Airbnb listing plotted on ISTAT areas



Finally, for each ISTAT polygon I added the information on Airbnb, by summing the listings, the amount of reviews, and the capacity (total number of beds, which can be multiple for each listing, per area).
This procedure was repeated for each of the seven cities, and I eventually merged the results into a single dataset.

4.2 Model

The model used for my regression is presented the following equation:

$$\ln Y_{ic} = \beta_0 + \beta_1 \ln(\text{Airbnb Listings})_{ic} + \beta_2 \text{ Population}_{ic} + \beta_3 \text{ Unemployment}_{ic} + \beta_4 \text{ Vacancy rate}_{ic} + \beta_5 \text{ Rent Market Size}_{ic} + \beta_6 \text{ College Share}_{ic} + \beta_7 \text{ stock conditions}_{ic} + \beta_8 \text{ OMI ring}_{ic} + \upsilon_{ic}$$

Where Y_{ic} is the average short-term rental price, obtained from the OMI areas, for the area *i* in city *c*.
The dependent variable is the logarithm of average rental prices, and the independent variable of interest is the logarithm of Airbnb listings. I will also make use of the count of Airbnb spots in levels, and the average rent in levels, by means of Negative Binomial Regressions.
Following the approach of Lee (2016),[40] I also create two more variables for measuring Airbnb's presence:

$$\textbf{Airbnb Density}_i = \text{Number of Airbnb Listings}_i \,/\, \text{Residential Housing Stock}_i$$
$$\textbf{Airbnb Capacity}_i = \ln(\text{Number of Airbnb Beds})_i$$

In order to isolate the effect of Airbnb on the STR rates of the micro-areas, I control for the factors that can influence this rates. All the other regressors are part of the ISTAT's population and households dataset. ISTAT census data was produced in 2011, and it is updated every 10 years (although it will be updated continuously with a sampling methodology, starting from the fall of 2018).[41]

The set of regressors are:

- Population
- Unemployment
- Rent Market Size: the number of households living in houses with a rental contract
- College Share: share of resident population with a bachelor degree or a higher

level of education

- Stock conditions: stock of residential buildings present in the area in terms of their qualitative characteristics over two categories, high quality and low quality.
- OMI ring / OMI belt: I constructed these two variables both in a progressive index from 1 to 5 (OMIring) and in dummies (OMIbelt), to represent the different OMI clusters in the city as in Figure 5 (central, semi-central, periphery, suburban area and rural areas)

Average STR rates are correlated with the median income of the resident population, and Barron et al. (2017) control for this factor in the model. Unfortunately, there is no data available at level of census tract to control for income. Plausibly, however, the stock conditions, college share, and the OMIring regressors can be intended as proxies for the different areas of the city, as will be further discussed during the interpretation of the results. Generally, I attempt to control for the single area effects, and isolate the association that Airbnb's presence has on rental prices. In fact, the cross-sectional nature of my dataset does not allow me to use a fixed-effects model as for Barron et al. (2017).

4.3 Descriptive Statistics

Table A summarizes the number of observations in my dataset, by city, where each observation corresponds to an ISTAT census polygon. Rome and Milan show the largest number of areas, cumulatively amounting to more than 60% of the 30806 observations.

Table A: Number of observations

City	**Frequency**	**Percent**	**Cumulative Percent**
Bologna	2037	6.61	6.61
Firenze	2066	6.71	13.32
Milano	5991	19.45	32.77
Palermo	2813	9.13	41.9
Roma	12604	40.91	82.81
Torino	3713	12.05	94.86
Venezia	1582	5.14	100
Total	30806	100	

Table B plots the percentage of Airbnb listings present in the cities by OMI ring. The last column highlights the share of Airbnbs concentrated in Central and Semi-central areas, confirming the expectation that a large share of Airbnb's is in the historic center, with a peak of almost 70% in Florence. Venice's lack in data depends on the OMI dataset, as there are several missing values for average prices for housing of residential and normal conditions.

Table B: Count of Airbnb listings by OMI ring

City	Central	Semi-Central	Periphery	Suburban	Rural	Central and Semi Central
Bologna	5.65%	12.96%	74.52%	6.87%		**18.61%**
Firenze	24.35%	44.58%	28.85%	2.23%		**68.93%**
Milano	12.65%	19.58%	59.11%	8.66%		**32.23%**
Palermo	35.02%	30.68%	15.29%	19.02%		**65.69%**
Roma	8.71%	21.35%	26.24%	43.39%	0.31%	**30.06%**
Torino	10.93%	37.11%	49.45%	2.50%		**48.05%**
Venezia	48.55%			51.45%		**48.55%**
Total	**15.04%**	**23.66%**	**36.45%**	**24.72%**	**0.13%**	**38.70%**

Table C depicts descriptive statistics for average measures of both ISTAT and OMI data. While Palermo is the cheapest city in terms of STR rates, Rome is the most expensive. When looking at average maximum prices, both Florence and Venice show peaking levels.

Table C: Summary statistics - ISTAT and OMI data

City	Population	Avg Rent (€/m^2 per month)	Avg sale price (€ / m^2)	Avg college Share	Residential Buildings	Avg Minimum Rent	Avg Maximum Rent
Bologna	356033	8	2399.9	39.9	20882	6.7	9.4
Firenze	358075	9.5	2849.3	34.8	31069	8.2	10.7
Milano	1238906	9.5	2941.7	44.8	42827	8.4	10.6
Palermo	650261	4.3	1382.3	28.1	44896	3.6	5
Roma	2613719	11.2	3093	39.7	136948	9.3	13.1
Torino	872367	6.8	1837.1	34.8	36158	5.5	8.1
Venezia	127123	9.5	2871.1	14.6	14526	7.7	11.3
Total	6216484	9.3	2682.4	37.4	327306	7.8	9.7

Table D summarizes measures of Airbnb's presence by city. In absolute terms, Rome is by far the city with the highest number of Airbnb listings, almost doubling those present in Milan. The third column contains a measure of average Airbnb usage, calculated as the number of average reviews per listing. This measure can be interpreted as the "intensity" of Airbnb utilization. Rome's score is again higher than that of Milan and the overall average across the cities is of 32.3 reviews per listing. Not surprisingly, cities like Venice and Florence, which are typically tourism-intense score on average 15 reviews more on every Airbnb. The average capacity of listings is similar for the seven cities, and is 3.4 overall, showing a high-presence of "entire-property" types.

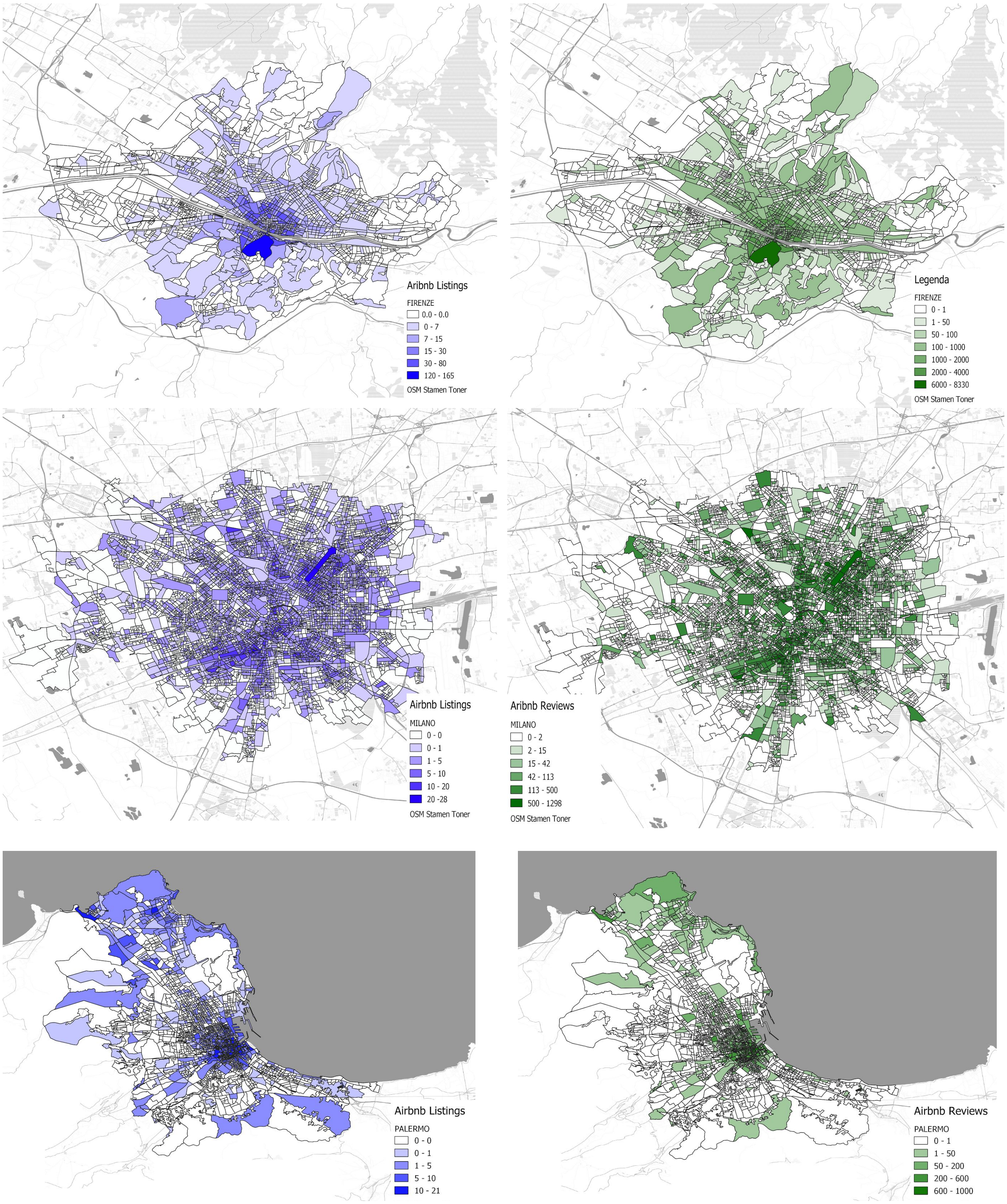
Airbnb density, instead, is defined as the number of active listings as a share of residential buildings present in the area. The last columns shows the number of listings per 1000 residents. Venice and Florence are once again scoring the highest values by far.

Table D: Summary statistics for Airbnb data

City	**Airbnb Listings**	**Total Reviews**	**Average Use (reviews per listing)**	**Total number of beds**	**Average Capacity**	**Airbnb listings per 1000 residents**
Bologna	3246	70217	21.63	10048	3.1	9.1
Firenze	6404	301563	47.09	22558	3.5	17.9
Milano	8525	235124	27.58	25993	3	6.9
Palermo	2672	55693	20.84	9985	3.7	4.1
Roma	16957	525561	30.99	58487	3.4	6.5
Torino	3088	66554	21.55	10010	3.2	3.5
Venezia	4209	200696	47.68	15619	3.7	33.1
Total	45101	1455408	32.27	152700	3.4	7.3

The same measures can be plotted on maps, in order to perform a visual analysis. The following figure illustrates the supply and demand for Airbnb in Florence, Milan and Palermo, where supply is proxied by the number of Airbnb listings, and demand is proxied by the number of reviews.

Figure 9: Supply proxied by listings (left) and demand proxied by reviews (right) for Airbnb, in Florence, Milan and Palermo

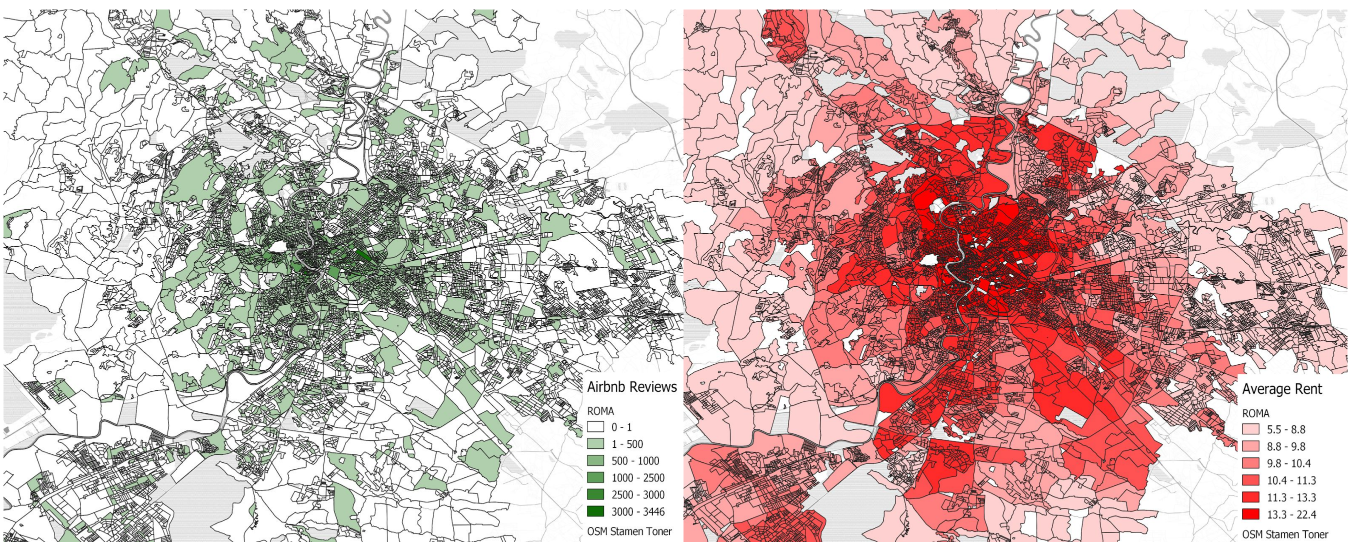


The reader can find similar maps for the other four cities in Figure 10 of the Appendix. Airbnb's presence is clustered in the historic centers of the cities, as it is especially visible in Florence and Palermo. Looking at Milan in Figure 9, is interesting to notice two sub-clusters, North-East and South-West of the very centre. They correspond, respectively, to two particular points of interest: the area immediately adjacent to Central Station, and the nightlife-renowned Navigli.

In Palermo it is possible to notice how demand and supply for Airbnb is also present in the northern coast just outside the city center, obviously associated with seaside-tourism and nightlife.

In general, the visual inspection can confirm the trend of Airbnb's presence highly associated with historic centers and tourism hotspots.

Figure 11: Demand and Average Rent in the center of Rome



Predictably, a higher demand and supply of Airbnb, is closely associated with higher average STR rates, as these dimensions are highly influenced by their relative centrality, as shown in Figure 11 for the centre of Rome. Figure 12 of the Appendix contains similar maps for the other six cities.
Having laid out a general overview of the dataset, in next Section I will proceed to present in detail the empirical strategy and the regressions results.

4.4 Regressions

The first regressions results presented in Table 1 tests with a Pooled Ordinary Least Squares regression the three measures for Airbnb's presence (as defined in Section 4.2). The dependent variable is the natural logarithm of the average rent, and the other regressors are expressed in logarithms too. All the beta coefficients are significant and positive.

Table 1: Pooled OLS, Standard Errors clustered by city

	(1)	(2)	(3)	(4)	(5)	(6)
	rent	rent	rent	rent	rent	rent
airbnb	0.101***	0.107***				
	(7.75)	(8.16)				
airbnb_dens			0.192***	0.199***		
			(6.98)	(8.11)		
airbnb_beds					0.0575***	0.0604***
					(9.49)	(12.75)
population	-0.0454	-0.0361	-0.0575	-0.0558	-0.0434	-0.0344
	(-1.04)	(-0.99)	(-0.96)	(-0.98)	(-1.03)	(-0.99)
collegeshare	0.158***	0.150***	0.164***	0.157***	0.155***	0.147***
	(4.44)	(5.48)	(4.41)	(5.67)	(4.11)	(5.05)
unemployment	-0.0681	-0.0664	-0.0623	-0.0615	-0.0681	-0.0665
	(-1.31)	(-1.36)	(-1.05)	(-1.10)	(-1.31)	(-1.38)
vacancy	-0.0122	-0.00112	-0.00473	-0.00169	-0.0117	-0.00104
	(-0.36)	(-0.03)	(-0.14)	(-0.05)	(-0.35)	(-0.03)
rentmarktsize	-0.0476**	-0.0460**	-0.0421*	-0.0391*	-0.0480**	-0.0464**
	(-2.74)	(-2.73)	(-2.03)	(-1.96)	(-2.82)	(-2.78)
stock_hq		-0.0117		0.0119		-0.0112
		(-0.40)		(0.41)		(-0.38)
stock_lq		-0.0517		-0.0414		-0.0498
		(-1.26)		(-0.96)		(-1.20)
N	30806	30806	27558	27558	30806	30806
adj. R^2	0.21	0.21	0.20	0.21	0.20	0.21

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The standard errors are clustered by city as the markets are obviously very different from each other, and Breusch-Pagan tests show signs of heteroskedasticity. As already mentioned, in order to control for the single microarea effect a proxy for income would be needed. It could be argued that the amount of graduates in the area are associated with income, through a standard returns-to-schooling effect. For the same reason, in columns (2), (4) and (6), I add the two variables representing the stock by quality offered per zone (where "*hq*" stands for high-quality and "*lq*" stands for low quality). As shown in the correlation table below, the correlation between low-quality stock and rent is -0.16. However, the correlation between high-quality stock and rent is of -0.05, meaning

Table E: Correlations Matrix

Variables	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
[1] rent	1								
[2] airbnb	0.2	1							
[3] airbnb_dens	0.221	0.614	1						
[4] airbnb_beds	0.211	0.867	0.698	1					
[5] population	-0.051	0.176	-0.037	0.187	1				
[6] collegeshare	0.19	0.306	0.115	0.342	0.826	1			
[7] unemployment	-0.067	0.164	-0.038	0.176	0.974	0.801	1		
[8] vacancy	-0.013	0.199	-0.013	0.196	0.323	0.307	0.311	1	
[9] rentmarktsize	-0.114	0.221	0.035	0.232	0.776	0.626	0.775	0.295	1
[10] OMIring	-0.208	-0.353	-0.393	-0.401	0.024	-0.207	0.013	-0.076	-0.127

A strategy to control for the within-city effects would be add the OMI rings of Figure 5 as regressor that distinguish between central areas, semi-central, peripheries and suburban and rural areas.

As shown in Table 1.1 below, once the rings are controlled for (in this specification expressed with the OMIbelt dummies), the three measures of Airbnb are still positive but insignificant, even in absence of other controls.

Table 1.1: Pooled OLS, Standard Errors clustred by city

	(1)	(2)	(3)
	rent	rent	rent
airbnb	0.0900		
	(1.43)		
airbnb_dens		0.226	
		(1.54)	
airbnb_beds			0.0541
			(1.56)
OMIbelt2	-0.108	-0.0787	-0.108
	(-0.72)	(-0.48)	(-0.72)
OMIbelt3	-0.229	-0.186	-0.227
	(-0.92)	(-0.71)	(-0.91)
OMIbelt4	-0.200	-0.153	-0.191
	(-0.85)	(-0.61)	(-0.81)
OMIbelt5	-0.196	-0.166	-0.193
	(-0.64)	(-0.55)	(-0.63)
N	30806	27558	30806
adj. R^2	0.11	0.08	0.11

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The dummy variable of reference is that for the central areas, and expectedly all the other coefficients are performing negatively with the respect to the omitted one.
The same happens when using the relevant variables in level, and exploiting the count data distribution of STR rates, by means of a negative binomial regression. The reader can find these estimates in Table 1.2 of the Appendix.

Table 2, instead, shows the OLS regressions for each individual city, with robust standard errors to correct for heteroskedasticity, after performing a Breusch-Pagan test.

Table 2: Pooled OLS by City

	(1) Bologna	(2) Firenze	(3) Milano	(4) Palermo	(5) Roma	(6) Torino	(7) Venezia
airbnb	0.0115**	0.0670***	0.0854***	-0.00522	0.0989***	0.0518***	0.0678***
	(2.52)	(21.03)	(10.78)	(-1.41)	(35.45)	(10.09)	(6.49)
population	-0.0502***	-0.0379***	-0.139***	-0.00669	-0.0915***	0.000151	-0.0473**
	(-6.04)	(-4.20)	(-9.23)	(-0.75)	(-16.47)	(0.02)	(-2.32)
collegeshare	0.0821***	0.0762***	0.231***	0.0761***	0.117***	0.0517***	0.0356***
	(15.39)	(19.17)	(39.32)	(34.34)	(51.46)	(22.07)	(3.95)
unemployment	-0.00524	-0.0510***	-0.0915***	-0.0275***	-0.0179***	-0.0579***	-0.0400**
	(-0.62)	(-5.71)	(-6.56)	(-3.18)	(-2.94)	(-8.25)	(-2.21)
vacancy	-0.00283	-0.000475	0.00894***	0.00573**	0.0425***	0.0105***	0.0760***
	(-0.71)	(-0.19)	(2.80)	(2.38)	(22.95)	(6.51)	(9.72)
rentmarktsize	-0.0151***	0.00655**	0.0355***	-0.00725**	0.0161***	0.00510**	0.00509
	(-3.67)	(2.27)	(8.94)	(-2.04)	(8.84)	(1.99)	(0.62)
stock_hq	0.0174***	-0.0130***	-0.0298***	-0.0284***	-0.0644***	-0.0185***	-0.0635***
	(4.32)	(-4.60)	(-6.16)	(-11.93)	(-27.79)	(-8.73)	(-8.00)
stock_lq	-0.0249***	0.00737**	-0.0624***	-0.0165***	-0.0159***	-0.00623***	0.0282**
	(-4.51)	(2.21)	(-8.76)	(-7.21)	(-6.82)	(-3.04)	(2.17)
N	2037	2066	5991	2813	12604	3713	1582
adj. R^2	0.18	0.49	0.31	0.35	0.41	0.27	0.33

t statistics in parentheses
Robust Standard Errors
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The coefficients are significant and positive for all the cities apart from Palermo, and the coefficients are similar, in terms of magnitude, similar to those found in the literature.

When the control for central, semi-central, periphery and suburban area is added, the effect of Airbnb is picked up by the stronger OMIring independent variable and becomes insignificant in Venice, while even negative, although extremely small, for Bologna, Palermo and Rome. The significant and positive effect is confirmed in Torino, Milan and Florence, at 1% confidence level, with R-squared values that are well above 0.5.

Table 2.1: Pooled OLS by City - full controls

	(1) Bologna	(2) Firenze	(3) Milano	(4) Palermo	(5) Roma	(6) Torino	(7) Venezia
airbnb	-0.0102**	0.0361***	0.0147***	-0.0150***	-0.00443**	0.0123***	-0.00481
	(-2.25)	(12.78)	(2.83)	(-3.92)	(-2.09)	(2.89)	(-0.52)
population	-0.0467***	-0.00731	-0.0413***	-0.0106	-0.000222	-0.00258	-0.0265*
	(-6.21)	(-1.01)	(-5.06)	(-1.24)	(-0.06)	(-0.52)	(-1.80)
collegeshare	0.0620***	0.0362***	0.0398***	0.0723***	0.0358***	0.0200***	0.0155**
	(11.51)	(9.40)	(10.09)	(33.40)	(20.74)	(10.57)	(2.03)
unemployment	0.0110	-0.0450***	0.00205	-0.0134	-0.0498***	-0.0203***	-0.00639
	(1.41)	(-6.46)	(0.27)	(-1.59)	(-12.86)	(-3.91)	(-0.46)
vacancy	-0.00662*	-0.000909	0.00171	-0.000409	0.00935***	-0.000101	0.0213***
	(-1.72)	(-0.40)	(0.83)	(-0.17)	(7.25)	(-0.08)	(2.75)
rentmarktsize	-0.0180***	-0.000550	0.00153	-0.0174***	0.00276**	-0.00805***	0.0121*
	(-4.53)	(-0.22)	(0.62)	(-4.77)	(2.11)	(-4.02)	(1.66)
stock_hq	0.0269***	-0.00259	-0.00959***	-0.0129***	0.00442***	0.00731***	-0.0154**
	(6.68)	(-1.01)	(-3.05)	(-4.69)	(2.63)	(3.84)	(-2.09)
stock_lq	-0.0204***	0.00867***	-0.0242***	-0.0107***	0.000282	0.00285	0.00177
	(-3.72)	(2.90)	(-4.73)	(-4.59)	(0.17)	(1.57)	(0.14)
OMIring	-0.0684***	-0.0861***	-0.343***	-0.0275***	-0.191***	-0.104***	-0.123***
	(-17.46)	(-24.60)	(-83.43)	(-10.58)	(-98.12)	(-34.99)	(-22.42)
N	2037	2066	5991	2813	12604	3713	1582
adj. R^2	0.24	0.62	0.73	0.38	0.73	0.54	0.52

t statistics in parentheses
Robust Standard Errors
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

However, I suspect that although the control for zones of the city might be useful in an attempt to isolate the effect of Airbnb in a cross-sectional dataset it tends to pick-up too much of the effect on rents. In fact, the OMI zones are built following a statistical methodology that tends to smooth out the average price per area, as explained in OMI's technical guide.[42]
The data on STR rates provided by OMI correlates with the rings at 0.26.

This is probably arising an issue of endogeneity in my estimates, which I cannot overcome with an instrument. Omitting this control, on the other hand, might cause and omitted variable bias that tends to overestimate the effect of Airbnb on STR rates.

A better approach could be that of clustering the standard errors by OMIring, instead of controlling for them: in such a way, the standard errors by city are clustered according to areas with similar characteristics, as shown in Table 2.2

Table 2.2: Pooled OLS, clustered by OMI ring

	(1) Bologna	(2) Firenze	(3) Milano	(4) Palermo	(5) Roma	(6) Torino	(7) Venezia
airbnb	0.0115	0.0670**	0.0854*	-0.00522	0.0989*	0.0518	0.0678
	(1.04)	(5.59)	(2.78)	(-0.48)	(2.35)	(1.26)	(4.00)
population	-0.0502***	-0.0379	-0.139**	-0.00669	-0.0915***	0.000151	-0.0473
	(-8.94)	(-1.69)	(-3.42)	(-0.43)	(-4.89)	(0.01)	(-1.62)
collegeshare	0.0821**	0.0762***	0.231	0.0761***	0.117**	0.0517*	0.0356
	(5.51)	(6.30)	(2.33)	(13.42)	(3.59)	(2.62)	(5.75)
unemployment	-0.00524	-0.0510	-0.0915	-0.0275	-0.0179	-0.0579*	-0.0400
	(-0.32)	(-1.25)	(-1.15)	(-1.59)	(-0.43)	(-2.69)	(-1.86)
vacancy	-0.00283	-0.000475	0.00894*	0.00573	0.0425	0.0105**	0.0760*
	(-0.31)	(-0.24)	(2.93)	(1.20)	(2.12)	(4.42)	(6.63)
rentmarktsize	-0.0151*	0.00655	0.0355*	-0.00725	0.0161	0.00510	0.00509
	(-3.13)	(0.53)	(3.11)	(-0.71)	(1.74)	(0.64)	(0.17)
stock_hq	0.0174*	-0.0130	-0.0298	-0.0284**	-0.0644*	-0.0185	-0.0635
	(2.39)	(-0.92)	(-1.89)	(-4.04)	(-2.35)	(-1.81)	(-5.07)
stock_lq	-0.0249***	0.00737	-0.0624	-0.0165	-0.0159**	-0.00623	0.0282
	(-7.48)	(2.10)	(-1.86)	(-2.31)	(-3.37)	(-0.62)	(0.76)
N	2037	2066	5991	2813	12604	3713	1582
adj. R^2	0.18	0.49	0.31	0.35	0.41	0.27	0.33

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The results remain significant and positive in Rome, Milan and particularly in Florence. However, the t-statistics reported in the parenthesis reflect an issue that once again arises from the cross-sectional nature of my dataset: in Venice, the standard errors are small, and the t-statistic is therefore as high as 4.00. However, I can cluster the standard errors only over two OMI rings, and one of them has too little observations, as OMI data contains several missing values.

The small number of clusters, in fact, affects the degrees of freedom, and therefore the p-values, and therefore yielding insignificant despite the high t-statistics. It is not possible even to calculate the Ibragimov-Muller cluster-adjusted t-statistics,[43] as not only the number of clusters is too small, but also the variation in rental rates is.
I will go back to this issue while discussing the limitations, and I will overcome it with my second empirical strategy (Section 5).

As a robustness test, I repeat the same approach in Table 2.3 of the Appendix using count data variables, by means of a negative binomial regression per city, confirming the positive and significant results in absence of the OMIring control.

The same regressions, by city, are performed with Airbnb density as the independent variable of interest.

Table 3 (in the Appendix) shows a strong positive association in absence of OMIring control between STR rates and Airbnb's density (measured as listings over number of buildings meant for residential use) in all the cities apart from Palermo, and a weak significance for Bologna.

When the control for OMIring is added, in Table 3.1, the results remain significant and positive for Florence, Rome and Torino. The coefficient for Florence is the most significant by far, as well as the coefficient reported in the homologous Table 2.1, showing a striking effect in a city with such an intense presence of tourists.

Table 3.1: Pooled OLS - Airbnb Density

	(1) Bologna	(2) Firenze	(3) Milano	(4) Palermo	(5) Roma	(6) Torino	(7) Venezia
airbnb_dens	-0.0183	0.122***	0.0144	-0.0241**	0.0314***	0.0340***	0.0785
	(-1.33)	(10.84)	(1.36)	(-2.23)	(5.10)	(3.61)	(1.35)
population	-0.0903***	-0.0476***	-0.0910***	0.00102	-0.0130***	-0.0273***	-0.109
	(-7.45)	(-5.61)	(-8.37)	(0.10)	(-3.03)	(-3.96)	(-1.12)
collegeshare	0.0742***	0.0491***	0.0616***	0.0726***	0.0406***	0.0238***	0.194***
	(12.13)	(11.72)	(14.73)	(32.08)	(21.99)	(11.93)	(4.13)
unemployment	0.0314***	-0.0219***	0.0240***	-0.0234**	-0.0447***	-0.00472	0.0127
	(3.44)	(-3.00)	(2.84)	(-2.48)	(-10.53)	(-0.75)	(0.14)
vacancy	-0.00768**	0.00179	0.00227	0.0000185	0.0100***	-0.000294	-0.00118
	(-1.96)	(0.79)	(1.12)	(0.01)	(7.65)	(-0.23)	(-0.04)
rentmarktsize	-0.0112**	0.00324	0.00879***	-0.0181***	0.00578***	-0.00495**	0.0107
	(-2.51)	(1.24)	(3.46)	(-4.57)	(4.17)	(-2.29)	(0.29)
stock_hq	0.0243***	0.00588**	-0.00782**	-0.0164***	0.00429**	0.00948***	-0.0891**
	(5.49)	(2.18)	(-2.49)	(-5.94)	(2.47)	(4.87)	(-1.97)
stock_lq	-0.0223***	0.0166***	-0.0211***	-0.0122***	-0.0000774	0.00486***	-0.181***
	(-4.06)	(5.76)	(-4.24)	(-5.12)	(-0.05)	(2.74)	(-2.74)
OMIring	-0.0612***	-0.0707***	-0.321***	-0.0248***	-0.179***	-0.0964***	-0.199***
	(-12.19)	(-20.04)	(-73.04)	(-9.19)	(-82.89)	(-30.76)	(-7.81)
N	1812	1836	5458	2592	11070	3408	1451
adj. R^2	0.25	0.59	0.74	0.40	0.71	0.53	0.13

t statistics in parentheses
Robust Standard Errors
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In Table 3.2 of the Appendix, I cluster again the Standard Errors by OMIring with Airbnb Density as the independent variable of interest finding significant outcomes again in Florence, Milan and Rome, although incurring in the same limitations expressed above.

4.5 Results and Limitations

The results reported in Section 4.4 present a positive and significant association between Airbnb's presence and the average STR rates, in the seven cities analyzed. The main limitation of this model is the cross-sectional nature of my dataset, that does not allow me to check the variation in Airbnb's presence and probably contributes to very small coefficients, which must be carefully interpreted. In Table 2, in absence of OMIring control, results seem to suggest that a 1% increase in Airbnb listings is associated with an increase of rental rates that ranges from the 0.01% of Bologna to the almost 0.1% of Milan and Rome.

I also try to control for the difference in areas within the same city, in lack of a good proxy for income. To fill this gap, further research might try build a similar instrument, or to stratify available open data at the ISTAT census tract level.[44]
When controlling for the within-city effects (central, semi-central, periphery, suburban and rural area) I encounter two main limitations. Firstly, an issue of endogeneity between the control and OMI's STR rates, discussed in detail in the previous section. Secondly, too little variation in prices in smaller cities like Venice, Palermo and Bologna, which compromises the significance of the results.

Although these coefficients are very small, we must take into consideration that STR rates are expressed in euros per square meters per month and that a cross-sectional model is not accounting for the time-varying effects, therefore ignoring of the explosion of the number of listings over the years, but only considering its atemporal presence in the neighbourhoods.

Generally, even at the cross-sectional level, my results seem to suggest that the presence of Airbnb is associated significantly with higher short-term rental rates in all the seven cities.
I find more robust results for Florence, Milan, Rome and Torino. The high significance of the results for Florence is in line with the theoretical expectations and the visual inspection, as the city is characterized by a high amount of tourists and a very concentrated presence of Airbnb.

In Section 5, I will overcome some of the limitations of this empirical approach, by using scraped data for STR rates.

5. Empirical Strategy II

5.1 Datasets

In order to obtain more granular and numerous observations on housing prices, I focused on the website Immobiliare.it,[45] a widely popular platform in Italy that lists properties for-sale and for-rent. The analogy for the United States would be Craigslist.[46]

The potential of this kind of detailed and real-time data is object of increasing attention by scholars. User-generated content is of high-interest in terms of research and policymaking: it can provide granular information about the stocks and flows of local real-estate markets in both time and space. Boeing and Waddell (2017)[47] scraped, cleaned an analyzed Craigslist as a mean to get new insights on the American housing market.

I decided to use a similar approach, and I built a web scraper to automate the collection of Immobiliare's listings for the rental market, in the seven cities considered in my analysis.

As Boeing and Waddell (2017) explain, there are obviously legal concerns in scraping data from these websites, which are addressed in their paper. I will mention two reasons why I decided to pursue this analysis. First, this research has a non-commercial nature: I have not sold, re-listed (neither for non-profit nor in a competing company), or re-published these information. In addition, I chose not to mask or change my IP address while performing the scraping, and neither to randomize the automatic consulting (which are common practices in web-scraping aimed at circumventing automatic blocking or blacklisting systems) and I was not banned or blocked from consulting the website.

In a very recent working paper by Bank of Italy, Loberto et al. (2018)[48] used Immobilare.it as a source for big housing data, explaining the high potential that they have for researchers. The paper also discusses in detail the cleaning procedure that would be necessary: the researchers implemented machine-learning algorithms in order to cut the several clones present, as many listings are re-posted.

My cleaning is, however, much more basic: I have deleted observations with the same identification codes (a value extracted during the collection) and afterwards those with the same URLs.

The dataset scraped and cleaned consists of 19325 observations for single properties. More than two thirds of the observations are in Rome and Milan.

The following Table summarizes the number of observations by City.

Table F: Distribution of observations by city

City	Frequency	Percent	Cumulative
Bologna	540	2.75	2.75
Firenze	865	4.41	7.16
Milano	4980	25.39	32.55
Palermo	1168	5.96	38.51
Roma	9373	47.79	86.3
Torino	2544	12.97	99.27
Venezia	143	0.73	100
Total	19613	100	

The data was scraped in May 2018. The STR rates listed on the website are generally higher than those of the OMI dataset, as these rats are offer price for the apartments.

5.2 Model

With Immobilare's dataset, I can construct a different model that exploits its extremely granular nature.

The model of reference for this empirical strategy is that of Linn (2013).[49]
Linn investigates the effects that sources of environmental contamination have on brownfields. Brownfields are real estate properties which value is influenced by the contamination: their renovation is hampered, and this, in turn, reflects negatively on the environment and welfare. The value of the properties can be influenced, he argues, both by known sources of contamination or by the *"concerns about associated liability"*.

The empirical analysis performed by Linn is based on hedonic regressions, where the independent variable of interest can influence the value of a single real estate property, while controlling for the characteristics of the property.

By the same reasoning, I treat Airbnb's presence as a market externality that is able to influence the price of single properties in the proximities.

The following equation shows the model used for the hedonic regressions:

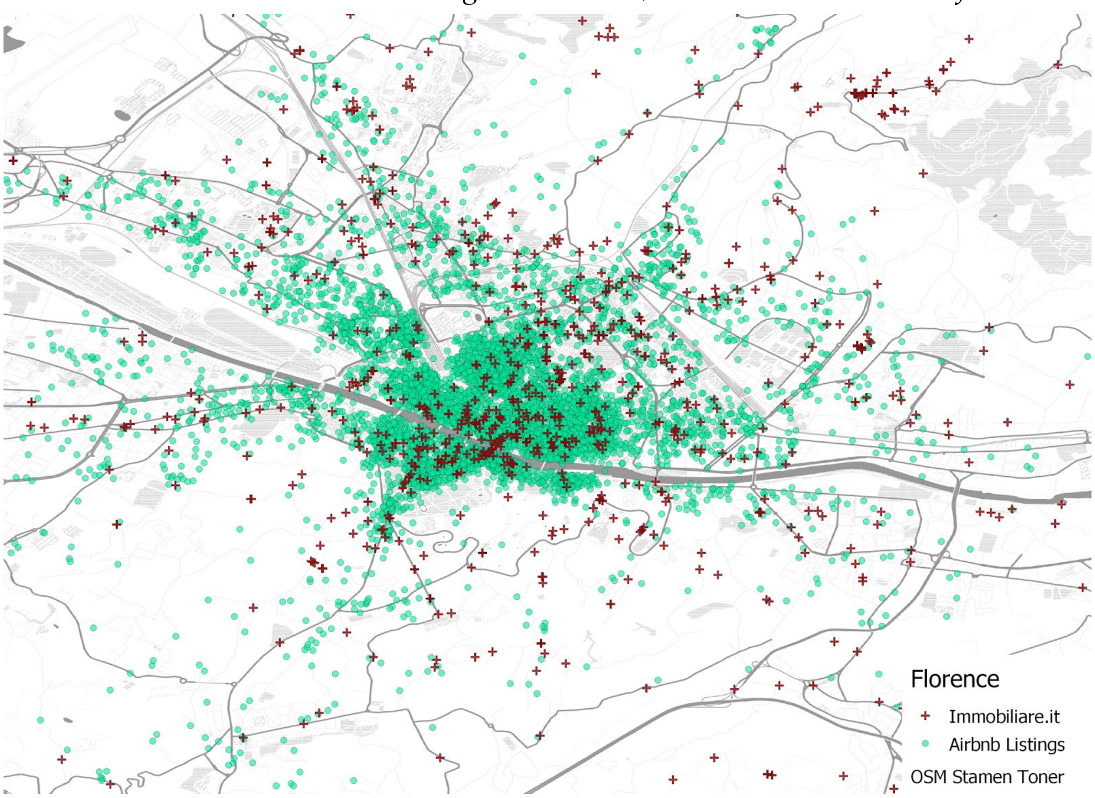
$$\ln Y_i = \beta_0 + \beta_1 \ln(\text{Airbnb Listings in proximity})_i + \beta_2 X'_i + \upsilon_i$$

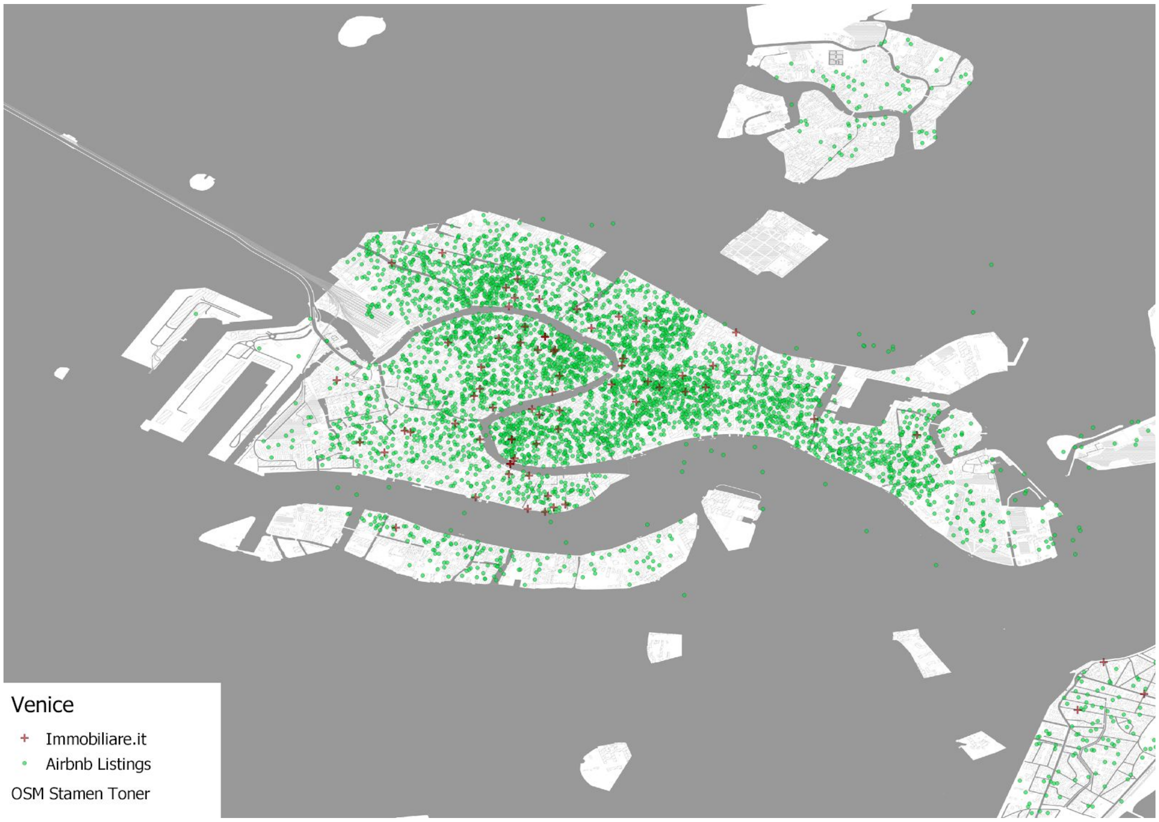
Where Y_i is the rental rate per square meter, and X'_i is a vector containing the characteristics of the property. During the data collection process, I was able to produce a vector containing:

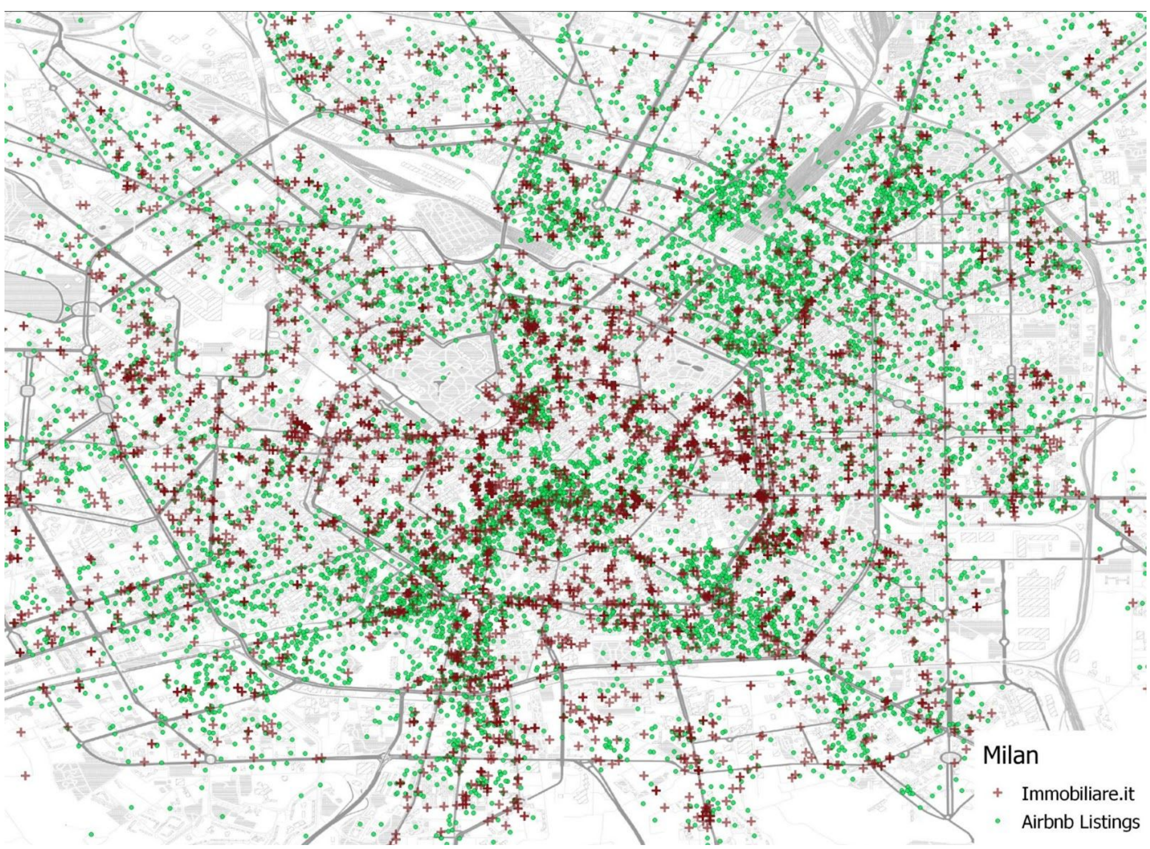
- Area of the property in square meters
- Number of rooms
- Number of toilets
- Property conditions expressed as 'optimal', 'good', 'mediocre' and 'bad'
- Year of construction of the building
- Energetic class

The following figures plot on map the Airbnb listings (green dots) and the listings scraped from Immobiliare (red crosses) in Florence, Venice and Milan.

Figure 13: Immobilare and Airbnb's listings in Florence, Venice and the center of Milan



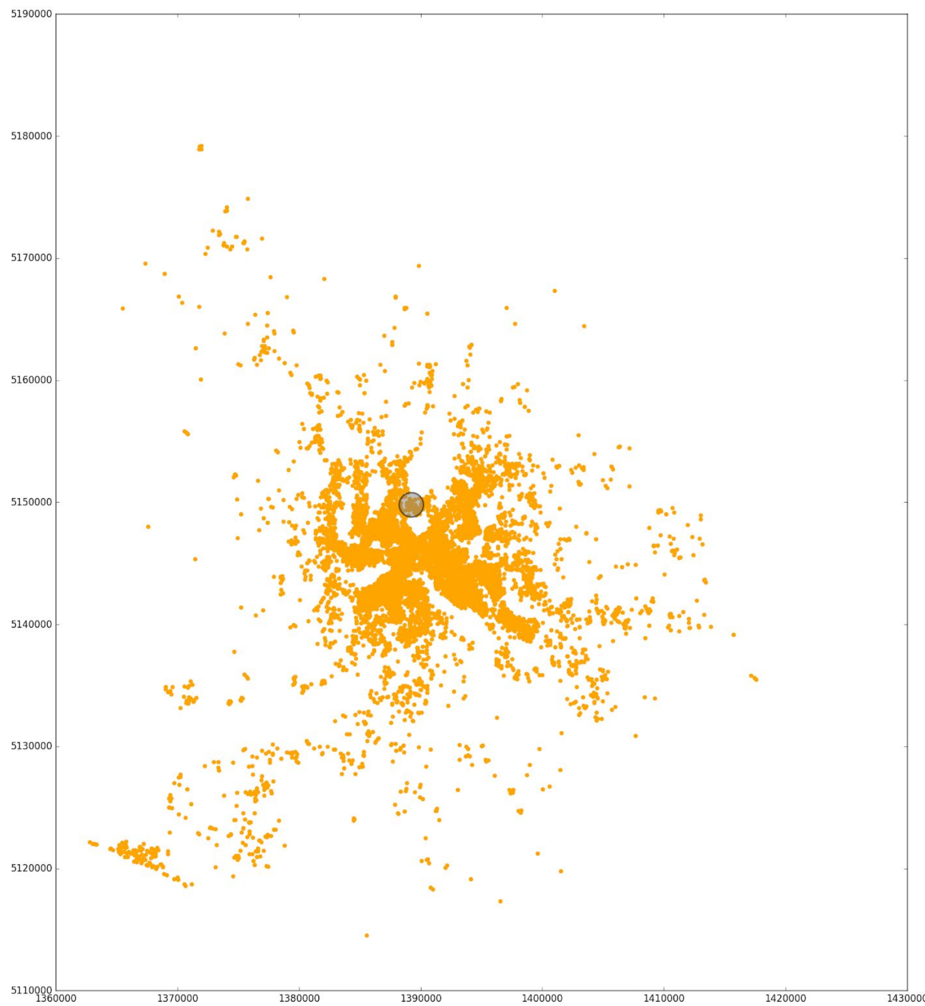

Venice
Immobiliare.it
Airbnb Listings
OSM Stamen Toner


Milan
Immobiliare.it
Airbnb Listings

Following the methodology of Linn (2013), I arbitrarily choose three radii to draw around the properties: 500, 750 and 1000 meters. I construct a count dataset by algorithmically counting the number of Airbnb's inside the radius, using the Python libraries GeoPandas[50] and Shapely.[51]

Figure 14 shows the buffering procedure, in which an apartment listed on Immobiliare is plotted, and after tracing a 1000 meters radius, the Airbnbs inside it are counted. I also merged this dataset with the previous one, by geolocating each observation inside an ISTAT and OMI area, in order to be able to use the previous controls.

Figure 14: Plotting and counting procedure



5.3 Descriptive Statistics

Table G presents the distribution of properties offered for rent by different OMI rings. More than half of the observations are listed in the central or semi-central areas, whereas 15% are located in the peripheries.

Table G: Distribution of Immobiliare Observations by OMI rings

OMIring	**Freq.**	**Percent**	**Cum.**
Central	4117	21.27	21.27
Semi-Central	6724	34.75	56.02
Periphery	5596	28.92	84.94
Suburban	2908	15.03	99.96
Rural	7	0.04	100
Total	19352	100	

It is also quite interesting to notice that, distributing the observations by energy efficiency index (where A is the most efficient and G the least), more than 80% of the stock offered is in the lowest two classes.

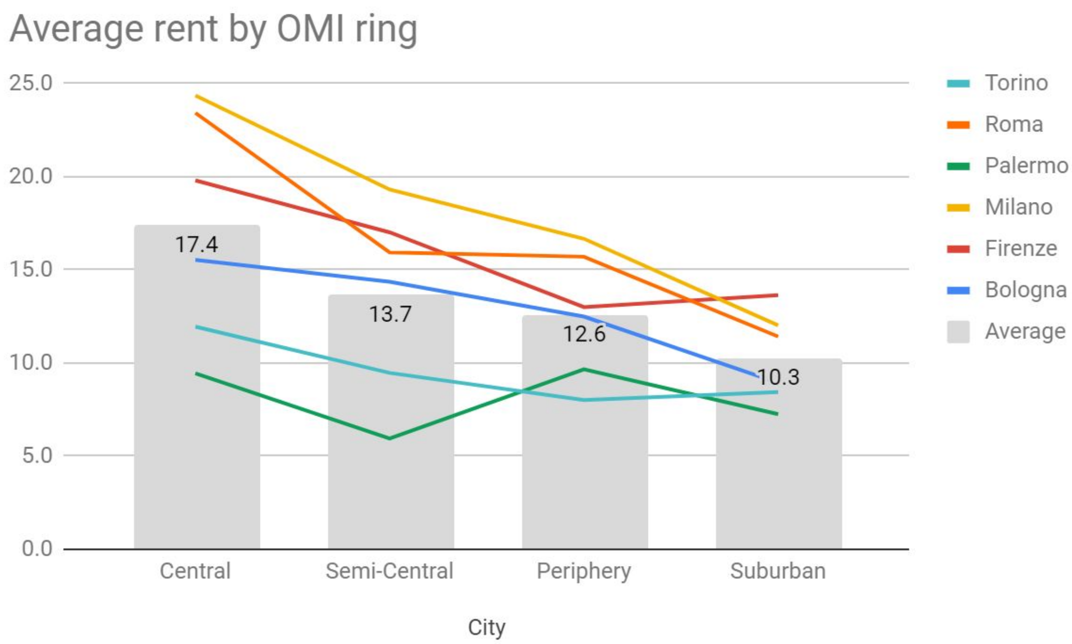
Table H: Distribution of Immobiliare's listings by energy efficiency

Energy Class	**Frequency**	**Percent**	**Cumulative percent**
A	433	2.63	2.63
B	319	1.94	4.57
C	435	2.64	7.21
D	826	5.02	12.23
E	980	5.95	18.18
F	1744	10.59	28.77
G	11727	71.23	100
Total	16464	100	

The graph in Figure 15 below, instead, presents an overview of the average prices, as they decrease by moving away from the centers. The graph excludes Venice, because of the lack of observations discussed above. The grey columns (labeled in black) show the average across the six cities.

Palermo and Bologna seem to have very little variation in prices across zones.

Figure 15: Average rental rate calculated on Immobiliare data, by OMIring



When looking at average rental rates, one must keep in mind that the prices listed on Immobliare's website are offer prices and the real value of the transactions might differ. What is of interest for my research, though, is their variation, so this does not represent a limitation *per se.*

In order to have a more precise insight on the differences in STR rates between the two sources, the reader can refer to Table J of the appendix, where it is possible to see the distance in prices between the OMI dataset and the Immobiliare dataset.

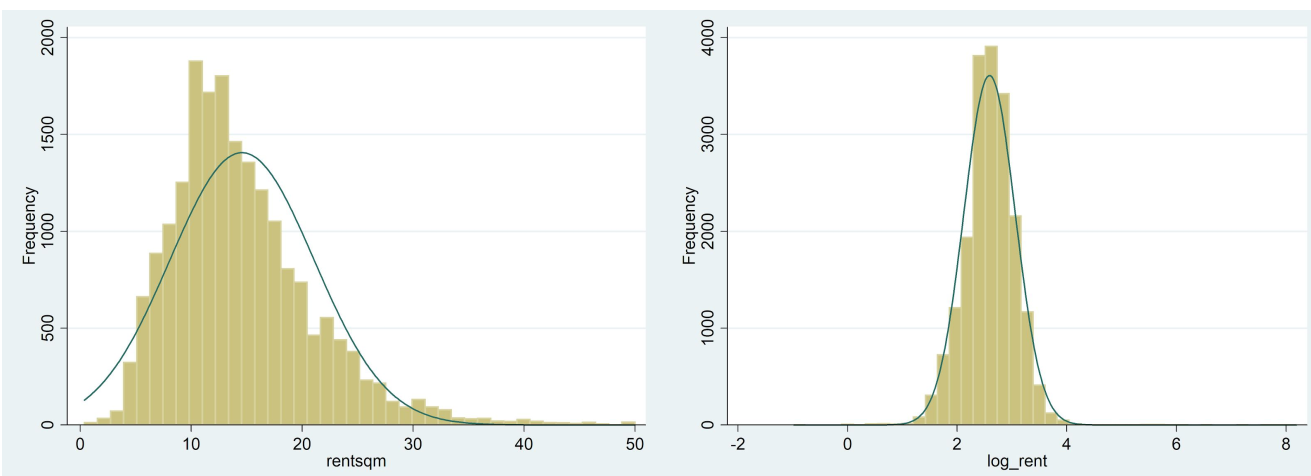
Table K presents information about the rental rates per city, as well as the count of Airbnb in the three radii drew around the properties offered for STR.
The highest average presence is in Venice, followed only by Florence. This is in line with the findings of Section 4.

Table K: Descriptive statistics for Airbnb's presence around Immobiliare's listings

City	Average Rent (€ per m²)	St. Dev. of Rent (€ per m²)	Average area (m²)	Avg Airbnb (500 m radius)	Avg Airbnb (750 m radius)	Avg Airbnb (1000 m radius)
Bologna	13.2	4.5	100.5	88.1	186.1	311.1
Firenze	17.4	16.9	112.9	195.1	413.6	695.7
Milano	19.8	35.8	101.8	58.0	123.1	210.0
Palermo	7.5	22.7	102.8	41.3	89.8	157.3
Roma	15.4	47.4	105.5	57.1	115.4	190.5
Torino	9.3	3.4	80.0	39.9	82.4	139.8
Venezia	14.6	13.4	96.5	211.4	438.6	708.6
Total	15.3	38.2	101.2	62.2	129.0	216.3

The following two figures plot the distribution of rents for my dataset, both in levels (excluding the outliers over 50€ per m^2 per month) and in the logarithmic transformation. I will exploit both these variables in the next Section, by means of pooled Ordinary Least Squares regressions.

Figure 16: Distribution of STR rates in level (€ per m^2 per month) and its logarithmic transformation



5.4 Regressions

In Table 4, I perform the hedonic regressions by controlling for the vector of property characteristics. The independent variables of interest are the logarithm of Airbnb's count inside the three radii.

For the sake of brevity, where the energy index and condition variables are marked with "yes" it means I am including all the dummies for these variables. Generally there are not big surprises on these factors, as the (few) properties with high energy class perform better than the lowest, and the same applies to property conditions. Once again, I perform the regressions clustering the standard errors by city.

Table 4: Pooled OLS, Hedonic Regression
Standard Errors Clustered by city

	(1)	(2)	(3)
	log_rent	log_rent	log_rent
lairbnb500	0.110^{***}		
	(4.76)		
lairbnb750		0.110^{***}	
		(4.47)	
lairbnb1000			0.110^{***}
			(4.31)
year	-1.320^{**}	-1.266^{**}	-1.246^{**}
	(-2.87)	(-2.71)	(-2.61)
Energy Index	Yes	Yes	Yes
Conditions	Yes	Yes	Yes
toilets	0.281^{***}	0.279^{***}	0.279^{***}
	(6.41)	(6.54)	(6.53)
rooms	-0.0869	-0.0877	-0.0889
	(-1.59)	(-1.63)	(-1.66)
ln_area	-0.222^{*}	-0.225^{*}	-0.226^{*}
	(-2.22)	(-2.31)	(-2.36)
_cons	13.08^{**}	12.61^{**}	12.40^{**}
	(3.44)	(3.26)	(3.14)
N	10046	10046	10046
adj. R^2	0.34	0.35	0.35

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All the result presented in Table 4 are positive and significant, with identical magnitude but decreasing significance, as the radius increases. Similar results are found using the variables in levels, as shown in Table 4.1 of the Appendix.
Given the similar coefficients but the higher significance of the 500 meters radius, I decide to show only results for this radius in the next tables. In regressions not shown I find almost identical results for the bigger radii.

In Table 5 I use pooled Ordinary Least Squares hedonic regressions for each city, with robust standard errors, given that Breusch-Pagan test showed again heteroskedasticity.
I also added the controls for population and share of population with a high level of education (*collegeshare*) obtained from the previous dataset, which obviously vary at a zone-level. The results are positive and significant in all the cities apart from Bologna and Palermo, confirming the results of Section 4.4.

Table 5: Pooled Methods - Hedonic Regression by city, robust standard errors

	(1) Bologna	(2) Firenze	(3) Milano	(4) Palermo	(5) Roma	(6) Torino	(7) Venezia
lairbnb500	-0.00620	0.0522***	0.0656***	-0.00118	0.0585***	0.0256***	0.127***
	(-0.44)	(5.25)	(10.10)	(-0.15)	(19.97)	(4.33)	(3.76)
lpopulation	-0.209***	-0.111***	-0.325***	-0.141***	-0.214***	-0.173***	0.0833
	(-4.23)	(-3.38)	(-17.91)	(-7.17)	(-22.28)	(-12.53)	(0.64)
lcollegeshare	0.224***	0.0944***	0.253***	0.131***	0.183***	0.159***	-0.000881
	(4.78)	(2.75)	(15.88)	(9.39)	(24.52)	(13.50)	(-0.01)
year	-0.501*	-0.370**	-2.289***	0.205	-1.900***	-0.852***	-3.440***
	(-1.74)	(-2.16)	(-7.22)	(0.28)	(-12.45)	(-2.67)	(-6.86)
Energy Index	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Conditions	Yes	Yes	Yes	Yes	Yes	Yes	Yes
toilets	0.0950**	0.138***	0.196***	0.0916***	0.149***	0.164***	0.270*
	(2.47)	(3.73)	(9.11)	(3.26)	(11.98)	(7.76)	(1.73)
rooms	0.00457	-0.0410**	0.0334*	0.0399	0.0233**	-0.0118	0.000206
	(0.12)	(-2.00)	(1.74)	(0.92)	(2.40)	(-0.87)	(0.00)
ln_area	-0.397***	-0.217***	-0.397***	-0.655***	-0.401***	-0.374***	-0.601***
	(-4.64)	(-3.77)	(-11.74)	(-4.78)	(-13.82)	(-12.07)	(-3.14)
_cons	8.349***	6.917***	22.11***	3.351	18.88***	10.32***	29.48***
	(3.70)	(5.14)	(9.26)	(0.60)	(16.19)	(4.28)	(7.69)
N	168	334	2639	735	4452	1192	36
adj. R^2	0.51	0.41	0.40	0.53	0.51	0.49	0.81

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In Table 6 I add to the regression the control for OMI rings (as already mentioned, expressed as a progressive variable for the sake of brevity, but I also tested it with the variable in dummies, finding almost identical estimates).

The effect of Airbnb on the single property STR rates is again positive and significant for Rome, Palermo, Milan and Florence.

Table 6: Pooled OLS - Hedonic Regressions by city

	(1) Bologna	(2) Firenze	(3) Milano	(4) Palermo	(5) Roma	(6) Torino	(7) Venezia
lairbnb500	0.00266	0.0248*	0.0244***	0.0201**	0.00796**	0.00388	0.0415
	(0.17)	(1.86)	(3.87)	(2.52)	(2.10)	(0.50)	(0.60)
OMIring	0.0347	-0.103***	-0.164***	0.0556***	-0.146***	-0.0520***	-0.149
	(0.96)	(-3.25)	(-17.60)	(3.09)	(-18.29)	(-3.75)	(-1.56)
lpopulation	-0.225***	-0.0575*	-0.164***	-0.154***	-0.144***	-0.161***	0.0579
	(-4.57)	(-1.67)	(-7.91)	(-7.04)	(-15.21)	(-11.51)	(0.42)
lcollegeshare	0.234***	0.0426	0.119***	0.141***	0.117***	0.152***	0.0460
	(4.98)	(1.26)	(7.03)	(8.87)	(15.30)	(12.88)	(0.35)
year	-0.594*	-0.361**	-1.548***	0.0881	-1.433***	-0.705**	-2.526***
	(-1.93)	(-2.17)	(-5.48)	(0.12)	(-10.26)	(-2.17)	(-2.86)
Energy Index	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Conditions	Yes	Yes	Yes	Yes	Yes	Yes	Yes
toilets	0.0998**	0.136***	0.176***	0.0885***	0.142***	0.164***	0.255
	(2.55)	(3.69)	(8.26)	(3.22)	(11.53)	(7.81)	(1.55)
rooms	-0.000666	-0.0386*	0.0404**	0.0425	0.0225**	-0.00909	0.0204
	(-0.02)	(-1.88)	(2.11)	(1.00)	(2.33)	(-0.67)	(0.41)
ln_area	-0.396***	-0.225***	-0.425***	-0.658***	-0.404***	-0.380***	-0.563**
	(-4.67)	(-3.83)	(-12.79)	(-4.93)	(-13.91)	(-12.31)	(-2.57)
_cons	9.132***	7.072***	16.79***	4.168	15.79***	9.356***	23.22***
	(3.88)	(5.51)	(7.90)	(0.73)	(14.85)	(3.81)	(3.55)
N	168	334	2639	735	4452	1192	36
adj. R^2	0.51	0.42	0.47	0.54	0.54	0.50	0.84

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In Table 7, I add both the average OMI rent (the rent calculated on the polygons of the previous Section) as well as the OMI rings variable to control for the within-city effects. Strikingly, I find significant results in Milan, Palermo, Rome, Torino and Venice, with R-squared values above 0.45.

The meaning of these last regressions is that Airbnb's presence in a 500 meters radius is significantly associated with higher STR prices of the properties, *ceteris paribus* the physical characteristics, the average rental rate of area and the area in which the property is (central, semi-central, periphery or suburban).

Table 7: Pooled OLS - Hedonic Regressions by city

	(1) Bologna	(2) Firenze	(3) Milano	(4) Palermo	(5) Roma	(6) Torino	(7) Venezia
lairbnb500	0.0226	-0.00764	0.0199***	0.0357***	0.00713*	0.0164**	0.126**
	(1.31)	(-0.59)	(3.21)	(3.93)	(1.93)	(2.02)	(2.33)
avg_rentOMI	0.245**	0.998***	0.364***	0.766***	0.842***	0.128	-0.128
	(2.20)	(5.67)	(14.42)	(10.51)	(24.71)	(1.36)	(-0.26)
OMIring	0.00569	-0.0691**	-0.0792***	0.0539***	-0.0420***	-0.0680***	-0.0563
	(0.16)	(-2.22)	(-7.24)	(3.06)	(-4.32)	(-3.64)	(-0.47)
year	-0.484	-0.122	-1.107***	0.143	-0.344***	-1.215***	-1.987**
	(-1.53)	(-0.73)	(-4.37)	(0.19)	(-3.07)	(-4.05)	(-2.34)
Energy Index	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Conditions	Yes	Yes	Yes	Yes	Yes	Yes	Yes
toilets	0.148***	0.0934**	0.166***	0.115***	0.135***	0.220***	0.294*
	(4.26)	(2.26)	(7.99)	(3.91)	(11.35)	(10.64)	(1.76)
rooms	-0.0209	-0.0348*	0.0416**	0.0310	0.0182*	-0.00766	0.0170
	(-0.53)	(-1.67)	(2.28)	(0.67)	(1.94)	(-0.53)	(0.35)
ln_area	-0.361***	-0.205***	-0.427***	-0.632***	-0.382***	-0.366***	-0.584**
	(-4.21)	(-3.30)	(-13.73)	(-4.36)	(-13.46)	(-11.12)	(-2.73)
_cons	7.222***	2.409	12.07***	1.893	4.701***	12.46***	19.10**
	(2.92)	(1.64)	(6.22)	(0.34)	(5.19)	(5.41)	(2.87)
N	175	348	2710	772	4670	1227	31
adj. R^2	0.45	0.44	0.49	0.51	0.58	0.44	0.82

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.5 Results and Limitations

The results of Section 5 confirm what Section 4 suggested: there is a positive and significant association between the presence of Airbnb and the short-term rent rates. Section 5 make use of extremely granular data on offer prices and single-property characteristics to perform hedonic regressions, treating Airbnb's nearby-presence as an externality. The results of the basic hedonic regressions are positive and significant. Even after controlling for single-neighbourhood effects, in its demographic and economic dimensions, I find significant coefficients. The results seem to suggest that a 1% increase in Airbnb's presence around 500 meters from a property can marginally increase its STR rate by 0.02% in Milan, 0.04% in Palermo, 0.016% in Torino and 0.13% in Venice.

As noted in Section 4.5, even if the coefficients of increase are very small, they refer to prices per square meter per month, and the evidence found at a cross-sectional level is ignoring the incredible expansion of Airbnb over time.
Even though using an extremely granular dataset allows me to find a residual effect of Airbnb even in a cross-sectional dataset, there are several limitations.

Scraped data is affected by three main types of errors.
Firstly, user-generated errors: I only roughly cleaned them by deleting evident outliers.
Secondly, as I explained in Section 3, random noise was applied at the source (as in the case of Airbnb), which might require time-consuming manual cleaning.
Third, in the case of Immobiliare.it I addressed the problem of duplicates in a very basic way, although machine learning techniques proved to be very effective.[52]

6. Conclusions

This research attempts to empirically understand the effect that Airbnb's presence has on the short-term rental rates in the seven Italian cities of Bologna, Florence, Milan, Palermo, Rome, Torino and Venice. After plotting the supply and the demand for Airbnb, I perform a visual analysis and present descriptive statistics, confirming the hypothesis that Airbnb is highly concentrated in the historic centers.

I make use of two different cross-sectional datasets and two different empirical strategies, in order to estimate the association between Airbnb's presence and STR rates. In the first strategy, I stratify data on prices and Airbnb's presence to the ISTAT census areas' level. In the second one, I perform hedonic regressions using highly granular scraped data. I find a significant and positive association in all the markets.

The association seems to be especially significant in Florence and Venice, two cities characterized by a strong tourism industry, a high intensity of Airbnb's usage and a particularly concentrated presence of listings in the city center. Previous attempts in the literature make use of panel data to isolate the effect of an Airbnb increase, finding positive results and coefficients of greater magnitude. In this sense, there is room for similar further research for the Italian case.

My research adds a piece to the increasing body of knowledge that tries to understand the price dynamics related to Airbnb in urban economies. There are, as mentioned, increasing concerns that home-sharing platforms are fostering the gentrification and touristization of historic centers, impacting spatial income inequality.

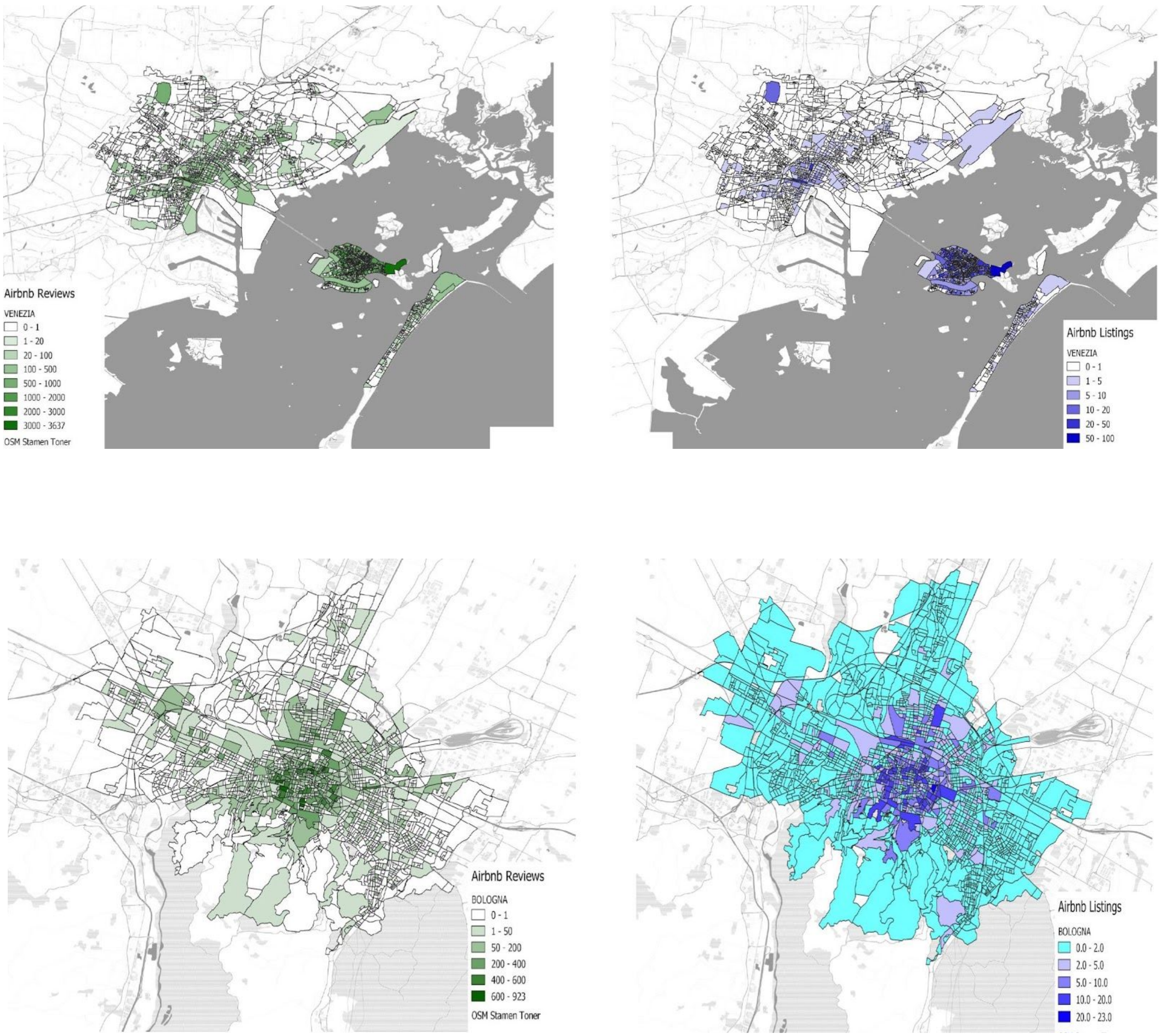
These concerns are being confronted by scholars, and policymakers are gradually addressing them by regulating the sharing economy. There are many ways to regulate home-sharing platforms: from price and booking caps, to taxes on short-term rentals. What is clear is that there is no one-fits-all solution, as the phenomenon is complex and somehow “silent”, also because of the voluntary lack of clean Airbnb data.

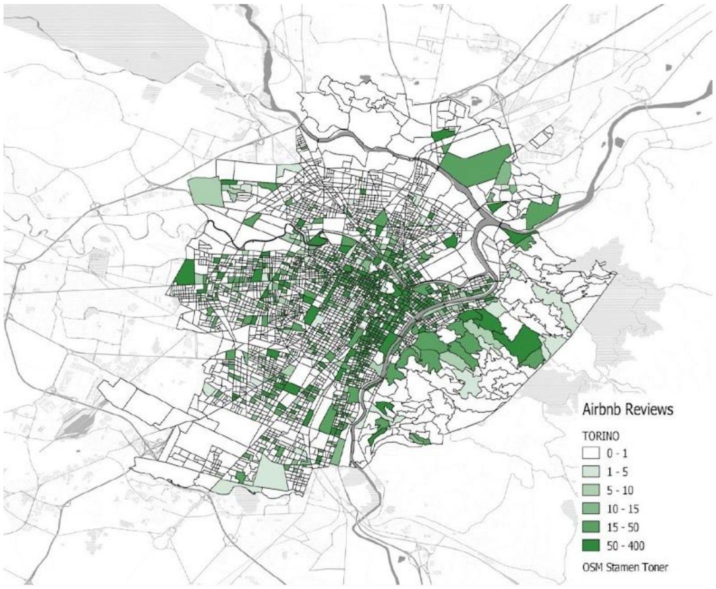
As explained by Picascia (2017)[53], Italy has implemented a simple flat tax rate of 21% on rentals shorter than one month. Although this is a first step, it might not be effective nor appropriate, given the nuanced and unequal nature of the issue.

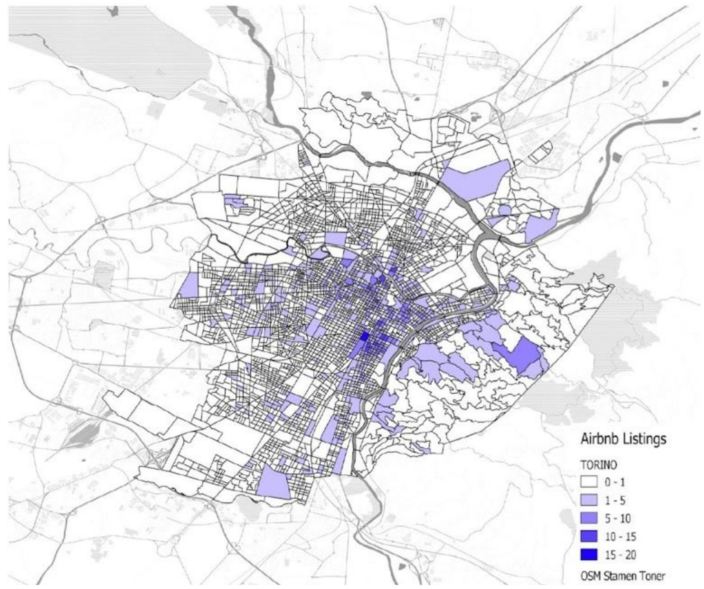
In my view, it is clear that we can no longer believe in the one-sided narrative of the sharing economy as a pure engine for cooperation, efficiency and innovation.
What we need, in the optic of protecting social welfare, is a serious, facts-based discussion about its broader impact on society.

Appendix: Figures and Tables

Figure 10: demand and supply (proxied respectively by reviews and listings) for Airbnb in Venice, Bologna, Torino and Rome:




Airbnb Reviews
TORINO
0 - 1
1 - 5
5 - 10
10 - 15
15 - 50
50 - 400
OSM Stamen Toner


Airbnb Listings
TORINO
0 - 1
1 - 5
5 - 10
10 - 15
15 - 20
OSM Stamen Toner


Airbnb Reviews
ROMA
0 - 1
1 - 500
500 - 1000
1000 - 2500
2500 - 3000
3000 - 3446
OSM Stamen Toner

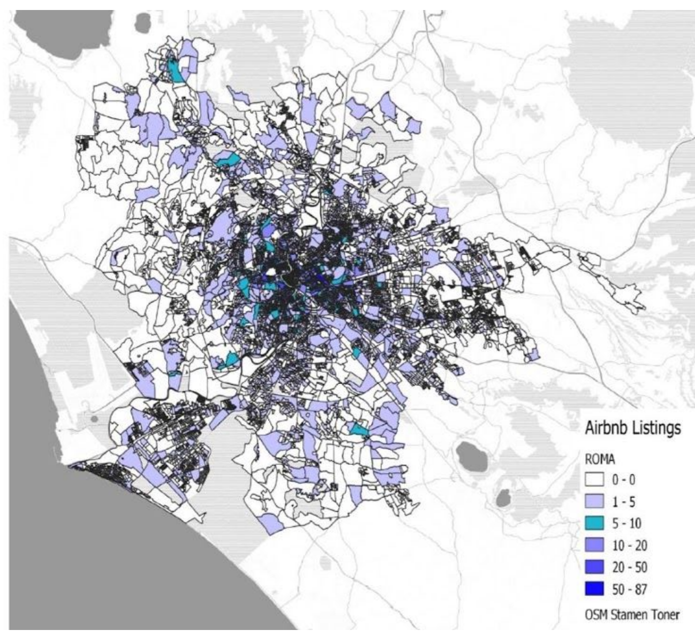
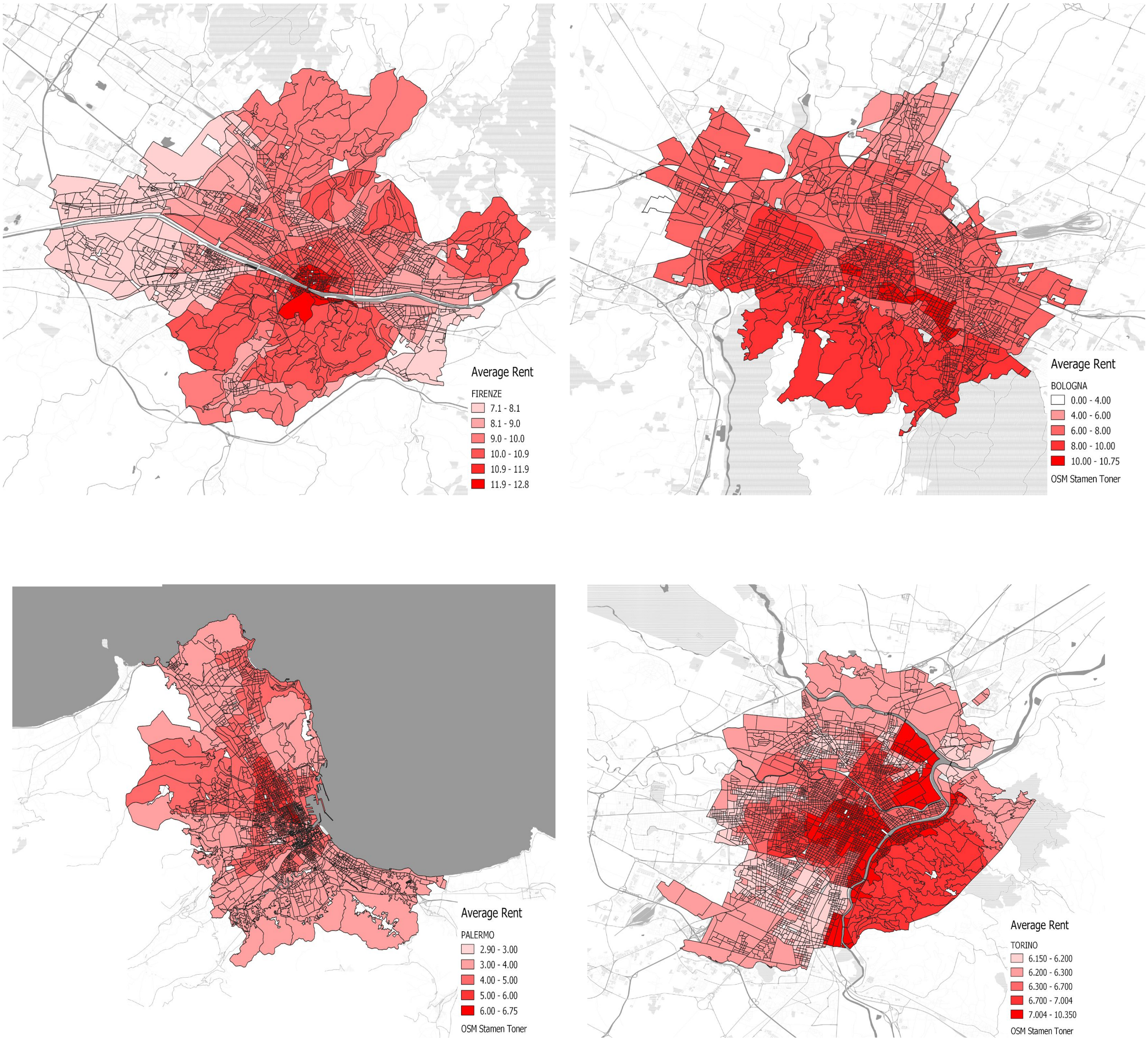
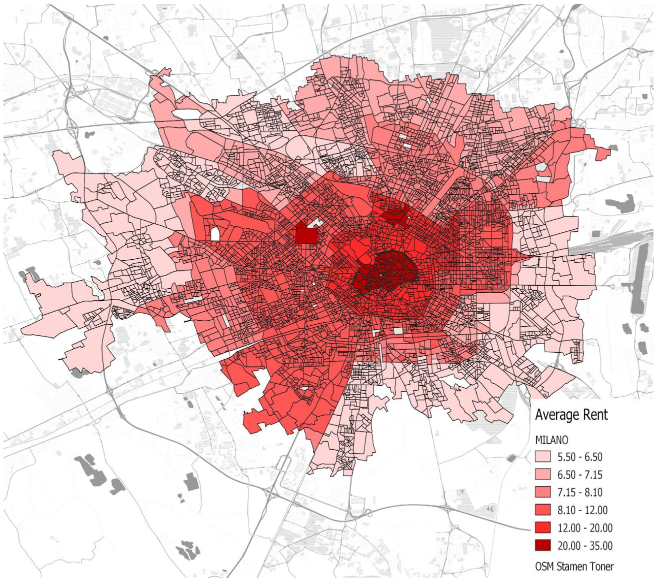

Airbnb Listings
ROMA
0 - 0
1 - 5
5 - 10
10 - 20
20 - 50
50 - 87
OSM Stamen Toner

Figure 12: maps of average short-term rental rates in Bologna, Florence, Milan, Torino and Venice





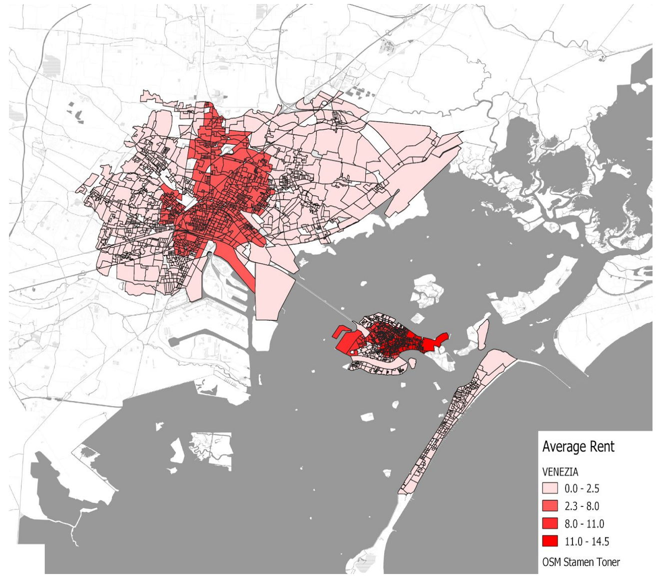


Table 1.2: Negative Binomial Regression
Standard Errors clustered by City

	(1)	(2)
	avg_rent	avg_rent
airbnb_count	0.0152***	0.00950
	(2.80)	(1.37)
population	-0.0612**	-0.0307***
	(-2.05)	(-2.90)
collegeshare	0.172***	0.131***
	(6.81)	(2.66)
unemployment	-0.0703	-0.0591
	(-1.53)	(-1.39)
vacancy	0.0143	0.00564
	(0.43)	(0.19)
rentmarktsize	-0.0385**	-0.0539***
	(-2.17)	(-3.68)
stock_hq	-0.0282	-0.00229
	(-0.88)	(-0.09)
stock_lq	-0.0391	-0.0409
	(-1.06)	(-1.08)
OMIring		-0.0917
		(-1.39)
N	30806	30806
adj. R^2		

t statistics in parentheses
Coefficients calculated as marginal effects at the mean
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.3: Negative Binomial Regression by City

	(1) Bologna	(2) Firenze	(3) Milano	(4) Palermo	(5) Roma	(6) Torino	(7) Venezia
airbnb_count	0.00315^{***}	0.00454^{***}	0.0264^{***}	-0.00142	0.0140^{***}	0.0174^{***}	0.00639^{***}
	(2.68)	(3.85)	(8.54)	(-1.49)	(13.23)	(10.52)	(3.37)
population	-0.0503^{***}	-0.0409^{***}	-0.152^{***}	-0.00806	-0.0965^{***}	0.000348	-0.0627^{***}
	(-6.04)	(-4.04)	(-8.25)	(-0.86)	(-15.17)	(0.05)	(-3.59)
collegeshare	0.0854^{***}	0.0984^{***}	0.268^{***}	0.0795^{***}	0.138^{***}	0.0557^{***}	0.0476^{***}
	(15.43)	(19.70)	(35.98)	(34.74)	(47.64)	(21.24)	(5.66)
unemployment	-0.00713	-0.0776^{***}	-0.126^{***}	-0.0280^{***}	-0.0359^{***}	-0.0628^{***}	-0.0422^{***}
	(-0.84)	(-7.39)	(-7.20)	(-3.08)	(-4.92)	(-8.18)	(-2.68)
vacancy	-0.00447	-0.00117	0.00932^{**}	0.00570^{**}	0.0538^{***}	0.0107^{***}	0.0958^{***}
	(-1.11)	(-0.43)	(2.40)	(2.32)	(24.31)	(5.77)	(14.32)
rentmarktsize	-0.0172^{***}	0.0202^{***}	0.0370^{***}	-0.00954^{***}	0.0189^{***}	0.00486^{*}	0.00452
	(-4.21)	(5.48)	(7.19)	(-2.59)	(8.53)	(1.67)	(0.65)
stock_hq	0.0199^{***}	-0.0123^{***}	-0.0347^{***}	-0.0290^{***}	-0.0786^{***}	-0.0197^{***}	-0.0601^{***}
	(4.83)	(-3.76)	(-5.69)	(-12.20)	(-29.29)	(-8.26)	(-8.69)
stock_lq	-0.0240^{***}	0.0156^{***}	-0.0731^{***}	-0.0161^{***}	-0.0153^{***}	-0.00801^{***}	0.0298^{***}
	(-4.28)	(4.07)	(-7.48)	(-6.99)	(-5.40)	(-3.50)	(2.86)
N	2037	2066	5991	2813	12604	3713	1582
adj. R^2							

t statistics in parentheses
Robust Standard Errors and coefficients calculated as marginal effects at the mean
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Pooled OLS by City - Airbnb Density

	(1) Bologna	(2) Firenze	(3) Milano	(4) Palermo	(5) Roma	(6) Torino	(7) Venezia
airbnb_dens	0.0239^{*}	0.192^{***}	0.129^{***}	-0.00381	0.234^{***}	0.105^{***}	0.394^{***}
	(1.82)	(13.96)	(7.37)	(-0.36)	(23.80)	(8.76)	(5.47)
population	-0.108^{***}	-0.0824^{***}	-0.294^{***}	0.000108	-0.121^{***}	-0.0333^{***}	-0.208^{**}
	(-8.30)	(-7.89)	(-13.35)	(0.01)	(-17.77)	(-3.55)	(-2.02)
collegeshare	0.0948^{***}	0.0861^{***}	0.264^{***}	0.0767^{***}	0.123^{***}	0.0553^{***}	0.245^{***}
	(16.27)	(20.58)	(41.55)	(33.37)	(49.99)	(22.10)	(5.05)
unemployment	0.0231^{**}	-0.0223^{**}	-0.00326	-0.0341^{***}	-0.00285	-0.0379^{***}	-0.00818
	(2.35)	(-2.50)	(-0.20)	(-3.50)	(-0.45)	(-4.56)	(-0.08)
vacancy	-0.00375	0.00298	0.0106^{***}	0.00534^{**}	0.0402^{***}	0.00911^{***}	0.0891^{***}
	(-0.94)	(1.24)	(3.50)	(2.16)	(21.81)	(5.81)	(3.72)
rentmarktsize	-0.00509	0.0110^{***}	0.0577^{***}	-0.00751^{**}	0.0237^{***}	0.0115^{***}	0.00698
	(-1.12)	(3.63)	(14.23)	(-1.99)	(12.33)	(4.29)	(0.18)
stock_hq	0.0217^{***}	0.00303	-0.0113^{**}	-0.0290^{***}	-0.0426^{***}	-0.00846^{***}	-0.116^{**}
	(4.83)	(1.03)	(-2.33)	(-11.59)	(-17.58)	(-3.93)	(-2.46)
stock_lq	-0.0231^{***}	0.0200^{***}	-0.0421^{***}	-0.0167^{***}	-0.00789^{***}	-0.000899	-0.109^{*}
	(-4.14)	(6.28)	(-6.21)	(-7.13)	(-3.40)	(-0.45)	(-1.73)
N	1812	1836	5458	2592	11070	3408	1451
adj. R^2	0.21	0.49	0.37	0.38	0.42	0.29	0.09

t statistics in parentheses
Robust Standard Errors
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.2: Pooled OLS, clustered by OMI ring

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Bologna	Firenze	Milano	Palermo	Roma	Torino	Venezia
airbnb_dens	0.0239	0.192^{**}	0.129^{**}	-0.00381	0.234^{*}	0.105	0.394
	(1.24)	(3.62)	(4.83)	(-0.16)	(2.18)	(1.45)	(1.83)
population	-0.108^{***}	-0.0824^{***}	-0.294^{*}	0.000108	-0.121^{**}	-0.0333	-0.208
	(-11.45)	(-9.24)	(-2.70)	(0.01)	(-3.06)	(-1.36)	(-4.36)
collegeshare	0.0948^{***}	0.0861^{***}	0.264^{*}	0.0767^{***}	0.123^{**}	0.0553^{*}	0.245
	(6.77)	(6.79)	(2.43)	(13.07)	(3.44)	(2.67)	(2.99)
unemployment	0.0231	-0.0223	-0.00326	-0.0341^{*}	-0.00285	-0.0379	-0.00818
	(1.07)	(-0.94)	(-0.09)	(-2.75)	(-0.12)	(-2.07)	(-0.08)
vacancy	-0.00375	0.00298^{*}	0.0106^{*}	0.00534	0.0402	0.00911^{**}	0.0891
	(-0.37)	(2.57)	(3.16)	(0.92)	(2.13)	(5.16)	(1.61)
rentmarktsize	-0.00509	0.0110	0.0577^{*}	-0.00751	0.0237	0.0115	0.00698
	(-1.21)	(1.09)	(2.75)	(-0.64)	(1.64)	(1.43)	(0.15)
stock_hq	0.0217^{**}	0.00303	-0.0113	-0.0290^{**}	-0.0426	-0.00846	-0.116
	(5.46)	(0.23)	(-1.02)	(-4.06)	(-2.12)	(-1.88)	(-2.80)
stock_lq	-0.0231^{***}	0.0200^{*}	-0.0421	-0.0167	-0.00789	-0.000899	-0.109
	(-6.13)	(2.92)	(-1.68)	(-2.20)	(-0.83)	(-0.11)	(-0.89)
N	1812	1836	5458	2592	11070	3408	1451
adj. R^2	0.21	0.49	0.37	0.38	0.42	0.29	0.09

t statistics in parentheses
Robust Standard Errors
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table J: Average price, by city and by OMI ring from OMI's dataset and Immobiliare.it dataset. Distance in prices = Immobiliare rate - OMI rate

OMI Rental Rates - Short-Term rental rates offers by OMIrings zones						
City	Central	Semi-Central	Periphery	Suburban	Rural	Total
Bologna	8.5	9.0	7.9	6.9		8.0
Firenze	11.3	9.1	8.6	7.8		9.5
Milano	17.5	11.4	7.6	6.1		9.5
Palermo	4.4	4.5	3.9	4.0		4.3
Roma	18.7	12.9	10.6	9.2	8.2	11.2
Torino	8.6	7.0	6.3	6.8		6.8
Venezia	11.6			7.5		9.5
Total	12.3	9.9	8.2	8.4	8.2	9.3
Immobiliare.it - Short-Term rental rates offers by OMIrings zones						
City	Central	Semi-Central	Periphery	Suburban	Rural	Total
Bologna	15.5	14.4	12.5	8.9		13.3
Firenze	19.8	17.0	13.0	13.6		17.7
Milano	24.4	19.3	16.7	12.0		20.0
Palermo	9.4	5.9	9.7	7.3		7.5
Roma	23.4	15.9	15.7	11.4	10.3	15.5
Torino	11.9	9.5	8.0	8.4		9.4
Venezia	21.9			10.4		15.8
Total	21.0	14.9	14.4	10.9	10.3	15.4
Distance in prices						
City	Central	Semi-Central	Periphery	Suburban	Rural	Total
Bologna	7.0	5.4	4.6	2.0		5.3
Firenze	8.5	7.9	4.4	5.8		8.2
Milano	6.9	7.9	9.0	6.0		10.6
Palermo	5.1	1.4	5.7	3.2		3.2
Roma	4.7	3.1	5.1	2.2	2.1	4.3
Torino	3.4	2.5	1.7	1.6		2.6
Venezia	10.3			2.9		6.3
Total	8.6	5.0	6.1	2.6	2.1	6.1

Table 4.1: Pooled OLS - Standard Errors Clustered by city			
	(1)	(2)	(3)
	rentsqm	rentsqm	rentsqm
airbnb500	0.0142***		
	(3.75)		
airbnb750		0.00758**	
		(3.56)	
airbnb1000			0.00490**
			(3.49)
build_year	-0.0200**	-0.0192**	-0.0187*
	(-2.48)	(-2.45)	(-2.40)
Energy Index	Yes	Yes	Yes
Conditions	Yes	Yes	Yes
toilets	3.074**	3.071**	3.066**
	(3.41)	(3.38)	(3.35)
rooms	-1.184**	-1.196**	-1.208**
	(-2.79)	(-2.78)	(-2.78)
area	-0.0115**	-0.0116**	-0.0116**
	(-3.11)	(-3.12)	(-3.12)
_cons	61.55**	60.10**	48.48**
	(3.36)	(3.32)	(2.88)
N	10046	10046	10046
adj. R^2	0.01	0.01	0.01

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Acknowledgements

This document represents the final chapter of my journey as an economics' student, and I would like to thank those who made the achievement of academic milestones possible.
The first, heartfelt, thank goes my parents Marinella and Marco and my brother Fabio. Absolutely nothing would be possible for me without the strength you give me every day: if I have any merit, that is also yours.
I want to thank Nadir Preziosi, my thesis supervisor, for his guidance during this year on both the academic and non-academic level.
I am deeply grateful to Christian Quintili for his friendship and mentorship throughout these years.
I sincerely thank Gianluigi Chiaro, Vincenzo Patruno, Andrea Borruso and Lorenzo Perone for their patience and their invaluable advices.
Last but not least, I want to thank Laura, Giovanni and Gaston, with whom I had the privilege to share a part of this journey: *ubi amici, ibidem opes*.

Notes and References

1. Rifkin, J. (2014). The zero marginal cost society: The internet of things, the collaborative commons, and the eclipse of capitalism. St. Martin's Press.

2. As appeared on Inc. Magazine: "Airbnb Is Inc.'s 2014 Company of the Year" - consulted on May 25th, 2018
https://www.inc.com/magazine/201412/burt-helm/airbnb-company-of-the-year-2014.html

3. As appeared on Business Insider - last consulted on May 25th, 2018
https://www.businessinsider.nl/airbnb-profit-revenue-2018-2/?international=true&r=US

4. "Airbnb: Fast Facts": https://press.atairbnb.com/fast-facts/ last consulted on May 25th, 2018

5. "Airbnb: About Us": https://press.atairbnb.com/about-us/ last consulted on May 25th, 2018

6. Pasquale, F. (2016). Two narratives of platform capitalism. Yale L. & Pol'y Rev., 35, 309.

7. The Guardian - "From Berlin to Barcelona; will Airbnb ruin our most loved cities?" - last consulted on May 28th, 2018
https://www.theguardian.com/technology/2016/jun/25/from-berlin-to-barcelona-will-airbnb-ruin-our-most-loved-cities

8. "Airbnb boom is 'emptying' historic Italian cities" - The Telegrph, last consulted May 28th, 2018
https://www.telegraph.co.uk/travel/destinations/europe/italy/articles/airbnb-boom-is-emptying-italian-cities/

9. "When Airbnb rentals turn into nuisance neighbours" - The Guardian, last consulted on May 28th, 2018

10. Pasquale, F. (2016), *op. cit.*

11. "Airbnb: Fast Facts", *op. cit*.

12. As appeared on ISTAT Press Release: "Tourism Satellite Account", consulted on May 28th, 2018 https://www.istat.it/en/archivio/208130

13. Real Estate Market Overview, 2017 - available online at
https://www.pwc.com/it/it/publications/assets/docs/pwc-real-estate-2017.pdf

14. Dagnes, J., Filandri, M., & Storti, L. (2018). Social class and wealth inequality in Italy over 20 years, 1993–2014. Journal of Modern Italian Studies, 23(2), 176-198.

15. Di Stefano, E. (2017). Leaving your mamma: why so late in Italy?. Review of Economics of the Household, 1-25.

16. Barron, K., Kung, E., & Proserpio, D. (2017). The Sharing Economy and Housing Affordability: Evidence from Airbnb.

17. Picascia, S., Romano, A., & Teobaldi, M. (2017, July). The airification of cities: making sense of the impact of peer to peer short term letting on urban functions and economy. In *Proceedings of the Annual Congress of the Association of European Schools of Planning, Lisbon* (pp. 11-14).

18. Lee, D. (2016). How Airbnb short-term rentals exacerbate Los Angeles's affordable housing crisis: Analysis and policy recommendations. Harv. L. & Pol'y Rev., 10, 229.

19. Quattrone, G., Proserpio, D., Quercia, D., Capra, L., & Musolesi, M. (2016, April). Who benefits from the sharing economy of Airbnb?. In Proceedings of the 25th international conference on world wide web (pp. 1385-1394). International World Wide Web Conferences Steering Committee.

20. Dudás, G., Boros, L., Kovalcsik, T., & Kovalcsik, B. (2017). The visualisation of the spatiality of Airbnb in Budapest using 3-band raster representation. Geographia Technica, 12(1), 23-30.

21. Horton, T. (2016). Reducing Affordability: The Impact of Airbnb on the Vacancy Rate and Affordability of the Toronto Rental Market.

22. Horn, K., & Merante, M. (2017). Is home sharing driving up rents? Evidence from Airbnb in Boston. Journal of Housing Economics, 38, 14-24.

23. Wachsmuth, D., Chaney, D., Kerrigan, D., Shillolo, A., & Basalaev-Binder, R. (2018). The high cost of short-term rentals in New York City. A report from the Urban Politics and Governance research group, School of Urban Planning, McGill University.

24. Cox, M. 2017. The Face of Airbnb, New York City: Airbnb as a Racial Gentrification Tool. Inside Airbnb.

25. Coles, P. A., Egesdal, M., Ellen, I. G., Li, X., & Sundararajan, A. (2017). Airbnb Usage Across New York City Neighborhoods: Geographic Patterns and Regulatory Implications.

26. Zervas, G., Proserpio, D., & Byers, J. (2014). The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry. Boston University School of Management Research Paper No. 2013-16.

27. Coyle, D., & Yeung, T. (2016). Understanding Airbnb in fourteen European cities. The Jean-Jacques Laffont Digital Chair Working Papers.

28. Edelman, B., Luca, M., & Svirsky, D. (2017). Racial discrimination in the sharing economy: Evidence from a field experiment. American Economic Journal: Applied Economics, 9(2), 1-22.

29. Picascia et al. (2017), *op. cit.*

30. Sampson, R. J. (2017). Urban sustainability in an age of enduring inequalities: Advancing theory and ecometrics for the 21st-century city. Proceedings of the National Academy of Sciences, 114(34), 8957-8962.

31. http://insideairbnb.com/behind.html

32. http://tomslee.net/

33. https://www.airdna.co/

34. https://github.com/tomslee/airbnb-data-collection

35. To whom I am extremely grateful

36. https://www.openpuglia.org/content/mission

37. "Association for the promotion of transparency and data culture through digital skills and investigative journalism" - http://ondata.it

38. Barron, Kung & Proserpio (2017), *op. cit.*

39. "Manuale della Banca Dati Quotazioni dell'Osservatorio del Mercato Immobiliare", Agenzia delle Entrate

40. Lee D. (2016), *op. cit.*

41. As appear on ISTAT website (last consulted on June 15th, 2018) https://www.istat.it/it/censimenti-permanenti/popolazione-e-abitazioni

42. "Manuale della Banca Dati Quotazioni dell'Osservatorio del Mercato Immobiliare" https://www.agenziaentrate.gov.it/wps/file/Nsilib/Nsi/Schede/FabbricatiTerreni/omi/Manuali+e+guide/omi+sintesi+manuale+bdq/Sintesi_Manuale_BDQ_OMI.pdf

43. Menger, A. (2017). CLUSTSE: Stata module to estimate the statistical significance of parameters when the data is clustered with a small number of clusters.

44. The cases would be possible, for example, using Bologna's Open Data at a statistical area level: http://dati.comune.bologna.it/node/244

45. https://www.immobiliare.it/

46. https://sfbay.craigslist.org/

47. Boeing, G., & Waddell, P. (2017). New insights into rental housing markets across the united states: web scraping and analyzing craigslist rental listings. Journal of Planning Education and Research, 37(4), 457-476.

48. Loberto, M., Luciani, A., & Pangallo, M. (2018). The potential of big housing data: an application to the Italian real-estate market.

49. Linn, J. (2013). The effect of voluntary brownfields programs on nearby property values: Evidence from Illinois. Journal of Urban Economics, 78, 1-18.

50. http://geopandas.org/

51. https://pypi.org/project/Shapely/

52. Loberto et al. (2018), *op. cit.*

53. Picascia et al. (2017), *op. cit.*