THE DYNAMICS OF REAL ESTATE PRICING

ORÇUN MORALI

BOĞAZİÇİ UNIVERSITY

2020

THE DYNAMICS OF REAL ESTATE PRICING

Thesis submitted to the

Institute for Graduate Studies in Social Sciences

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Management

by

Orçun Moralı

Boğaziçi University

2020

DECLARATION OF ORIGINALITY

I, Orçun Moralı, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature ....

Date ....22. 07. 2020..

ABSTRACT

The Dynamics of Real Estate Pricing

This study is a detailed examination of pricing dynamics in real estate markets, particularly in the housing market with a focus on the shortcomings of standard pricing models.

Two behavioral tendencies, even pricing and precision avoidance, are established in real estate markets. The tendencies are documented, quantified and shown to be an anomaly in real estate markets. The determinants of the tendencies are attributed to negotiation and cognition processes.

Dynamics of spatial dependence are examined through a spatial analysis of housing using geo-locations of properties. Determinants of spatial dependence are identified and quantified.

As a behavioral real estate study, this study also presents findings on to the generalizability of the theories in different markets, such as in different cultures or in developed or emerging markets.

ÖZET

Gayrimenkul Fiyatlaması Dinamikleri

Bu çalışma, gayrimenkul pazarlarındaki fiyat dinamiğini, standart fiyatlama modellerindeki eksiklikleri ağırlıkla konut pazarı üzerinden ele alarak detaylı olarak inceleyen bir çalışmadır.

Gayrimenkul pazarlarında iki davranışsal eğilim -yuvarlak fiyatlama ve hassas fiyatlamadan kaçınma- tespit edilmiştir. Bu eğilimler belgelenmiş, etkileri sayısallaştırılmış ve gayrimenkul pazarlarında bir anomali olarak değerlendirilmiştir. Bu eğilimlerin oluşmasında pazarlık süreçlerinin ve bilişsel süreçlerin etkili olabileceği belirlenmiştir.

Bu çalışmada ayrıca, konut pazarlarında mekansal bağımlılığın dinamikleri, konutların coğrafi koordinatları kullanılarak yapılan mekansal bir inceleme ile analiz edilmiştir. Mekansal bağımlılığın etkenleri belirlenmiş ve sayısallaştırılmıştır.

Bir davranışsal gayrimenkul çalışması olarak bu çalışmada, teorilerin farklı pazarlarda, örneğin farklı kültürler altında veya gelişmiş ve gelişmekte olan ekonomilerde, nerelerde genellenebildiğine ve genellenemediğine dair bulgular sunulmuştur.

ACKNOWLEDGEMENTS

This thesis was partially funded by Boğaziçi University Scientific Research Projects (Boğaziçi BAP) Fund with the projects numbered 12821 and 14562.

This thesis uses the tools provided by the open source community, both by the core developers and package developers, around The R Project for Statistical Computing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDIX TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND REAL ESTATE MARKETS IN TURKEY

1.1 Preface

Financial markets are a main theme for finance researchers. It consists of a diverse set of markets, including and not necessarily bound to equity markets, fixed-income securities markets, derivative markets, debt markets, commodity markets and real estate markets. Finance researchers are usually fortunate in terms of data, as financial data is plentiful, measured well, frequent and of high quality, compared to other social sciences. Financial markets is a core concept in the financial world, along with institutions, investors, firms, governance and household.

The discussions in finance, however, might exceed the financial world in their implication. The excessively discussed concept of market efficiency, for instance, has merged with the research in fields like psychology and economics to have implications on overall regulations, institution design, governance and policy decisions. As fascinating as it is to observe the discussions of finance researchers and the far implying reach of results, this dissertation is written to touch topics in areas far less seen and focused on nevertheless I find very significant.

Studying the literature from an international perspective in Turkey, it was easier to think about how market dynamics could be fundamentally different based on regulations or culture. Most financial research is conducted in developed economies, where the regulations are much more stable, market culture is more established and abrupt changes in market conditions are less frequent. In contrast, emerging economies have less stable market conditions and the interference of regulations or big actors is much more visible. Such differences in market characteristics between developed and emerging economies do not necessarily constitute a problem in financial research, but it is a call for testing before automatically applying interpretations from one market to another. In that regard, testing the research in emerging markets is crucial in judging the generalizability and robustness of general financial discussions. This thesis, in part, touches on merging the theories on developed markets and emerging market data.

The methodological advances in finance have been very quick in the sense that both the amount of data and available econometric and statistical methods and tools that can make use of the data have greatly increased in the last few decades. One of the econometric advances that happened after the 1950's and gained traction after 1980’s is spatial econometrics, that tries to correct calculations by incorporating a space definition. Geography is a fundamental factor that has an underlying effect on things in complex ways. My belief is that markets are not immune to this, and for that reason, spatial econometrics is to be added to the toolbox of financial research. This thesis touches on applying spatial econometrics techniques on real estate pricing.

A contemporary discussion in financial research is behavioral finance and what it means for market efficiency. Initially attracted a lot of backlash, behavioral approaches have managed to prove themselves to be highly respectable. There has been a great amount of literature accumulated on behavioral finance, regardless there are still parts of research needs to be done. Particularly, the vast majority of the research is based on equity markets, therefore extending the research to other asset classes would help in testing and generalizing implications from existing research. This thesis contributes to carrying behavioral approaches outside of equity markets to real estate markets. Behavioral research in real estate markets is still in its early stages.

In short, this study is relevant in current literature as it promotes three ideas. First, the economic environment and the dynamics of economics in emerging markets could be fundamentally different than developed markets and therefore automatic adopting of policies is not appropriate. Second, the effect of geography and the network by which the processes occur is hard to measure and therefore ignored in most research. Regardless, the use of spatial econometric techniques can help measuring effects of factors in better ways. Third, the weight of research in equity markets compared to its weight in overall markets are high and extending behavioral financial research to other asset classes than equity markets, particularly real estate markets, is needed.

1.2 Introduction

This dissertation is divided into 4 chapters, each of them being independent studies that can be read separately without any contextual problems, should the reader choose to do so. Yet, they all fall under the umbrella of real estate markets with a financial point of view and they complement each other in obtaining a comprehensive understanding of real estate markets, in perspectives, and in research methodologies and understanding limitations.

This first chapter is an introductory chapter with an aim to provide background information on the particular markets and domain to which data belongs. The next section of this chapter articulates the motivation for selecting real estate markets as the general field. Then, the fundamentals of the Istanbul metropolitan real estate market is explained with a particular focus on housing, in order to get the domain knowledge of the data used. Next, the data used in this thesis is described and some descriptive statistics are provided. More details regarding data, that are particular to research questions, are later explained in detail in their relevant chapters. After the data section, general methodological approaches and tools used are provided. Finally, a summary and policy implications are stated.

The second chapter presents a survey of spatial econometric pricing techniques, which are beneficial tool boxes in real estate research but yet missing in financial research. A discussion on the techniques for the cross-section is given and spatial terminologies used in this thesis are defined. This chapter is also relevant in understanding the discussions in Chapter 3.

The third chapter exhibits a spatial analysis in the housing market. It identifies and tests the determinants of spatial dependence in the housing market, using various spatial measures and econometric techniques. This chapter also quantifies some of the spatial effects, enabling a better understanding of the limitations of standard hedonic regressions.

The fourth chapter unveils a behavioral real estate study. In this chapter, the effects of behaviors even pricing and precision avoidance on prices are detected. A pricing anomaly from the even-pricing phenomena is identified and documented in multiple real estate markets. Possible explanation for this pricing behavior is examined and discussed within an understanding of market-efficiency.

1.3 Motivation

The real estate market is an important part of the economy, it is bigger in size than the well-studied equity markets and the market participation rates are higher in real estate. The total global real estate values totaled 217 trillion dollars, -around 75% being residential property- while the total of the global equity and outstanding securitized debt is worth $149trillion.[1] Moreover, the market participation rates are higher in real estate; the stock market participation is around 20% on average (Giannetti & Koskinen, 2010) compared to an average of 60% homeownership (Andrews & Sánchez, 2011). These comparative figures are even more striking in Turkish markets. The stock market participation is around 1.2% (Giannetti and Koskinen, 2010) whereas homeownership is 67.3%[2] in Turkey. Over the last two decades, the real estate activities on average accounted for around 12% of all sectoral activities in Turkish GDP.[3]

In addition to the depth and penetration of real estate markets, the purchasing of a real estate property is one of the most important financial decisions for most individuals as it includes a significant portion of the individual's wealth. Even one percent of the median house prices is an economically significant amount which is comparable to months' worth of disposable median income. Consequently, any inefficiencies in the market is of direct importance to the general well-being.

[1] What Price the World? Trends in International Real Estate Trading, Savills Research, Report of 2016 Global Real Estate Tips, January 2016.

[2] Turkstat Population and Housing census, 2011. http://www.tuik.gov.tr/PreHaberBultenleri.do?id=15843.

[3] TURKSTAT GDP statistics. http://tuik.gov.tr/PreIstatistikTablo.do?istab_id=2515. Real estate activities include buying, selling, renting and management of real estate, as defined by NACE Rev. 2.

Investment in real estate not only serves the need for shelter for most people, it is also a means to store wealth. It is estimated that around 60% of all wealth is stored in real estate markets globally.[4] Such figures are so striking that it makes one reconsider the relationship between money and real estate as to which is measured in which to what degree. Considering the importance of real estate markets, it follows as a natural choice for the general domain for financial research.

1.4 Fundamental overview of real estate markets in Turkey

In order to provide a context for the data used in this thesis and to understand the domain thoroughly, a fundamental look on the market is presented in this section.

The effect of macro trends are important in understanding the real estate market in Turkey. Turkey faced a strong globalization trend during the 1980's. During this period, Turkey has gradually moved from being a closed economy to an open one. The restrictions on imports are removed, the currency regime is changed to a floating regime and the economy became globally integrated. This period corresponds to a high-growth and high-urbanization period in Turkey real estate market. There is not much reliable public data about real estate prices in Turkey for these periods; however, it’s been common knowledge that it was a period of wealth transfer to those who are invested in Istanbul real estate market and to those who are settled in Istanbul without land permissions and are eventually allowed to claim the ownership of the land.

[4] What Price the World? Trends in International Real Estate Trading, Savills Research, Report of 2016 Global Real Estate Tips, January 2016.

The oldest house prices indices in Turkey goes back to 1880’s; however, the data scope is very limited for most of this period. The house price indices of financial research quality that are used commonly are provided by the Central Bank of Turkey and prices are available since 2010. Researchers need to keep in mind that this data does not allow much generalizations over time as it mostly corresponds to a boom period with historically very low interest rates with respect to Turkey’s economy. Figure 1 shows the hedonic house price index (HHPI), which is a nominal index corrected by the increased quality of housing.

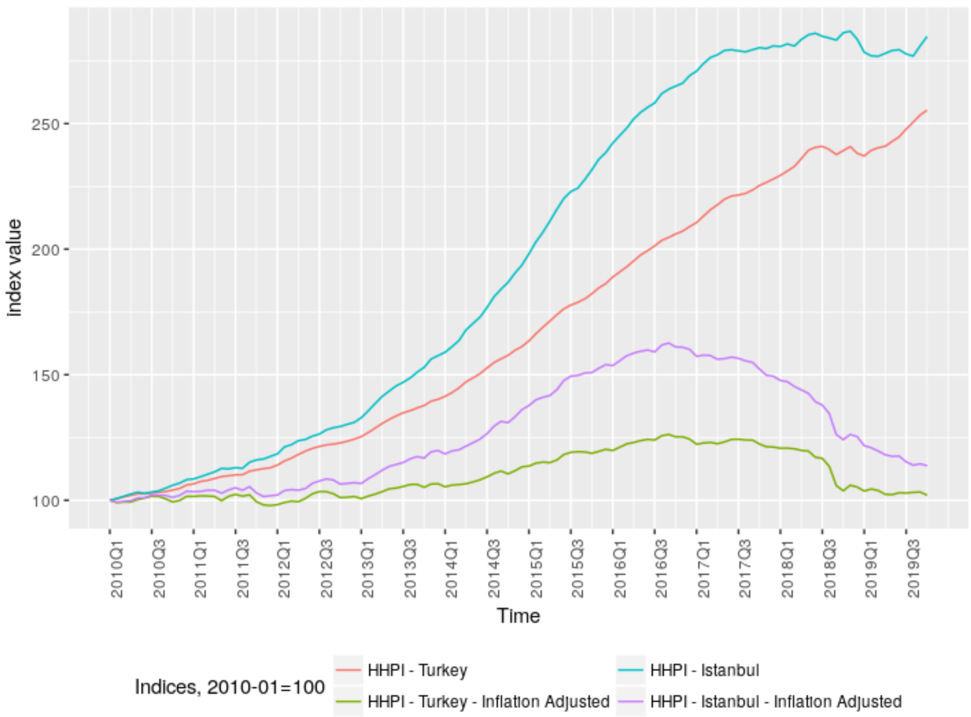

Figure 1 Nominal and real hedonic housing price indices for Turkey and Istanbul, 2010-01 = 100

Source: Author’s calculation based on CBRT data.

Another economic reality of Turkey significant to real estate markets is that it has been a high-inflation country historically. It has faced hyperinflation and inflation-related crises in its past. Such an environment is a cause of constant nominal increase in real estate prices, which puts the practitioners in a mindset that the listing prices are always to be set higher. This belief is not always true in real terms as easily seen in Figure 1. It is also generally believed among the financially literate that real estate is a good hedge against inflation, which goes along with the constant nominal increase in real estate prices. This is, again, not necessarily true overall for an arbitrary selection of period as seen in Figure 2, which compares real hedonic housing returns over the years to other investment opportunities since HHPI is available.

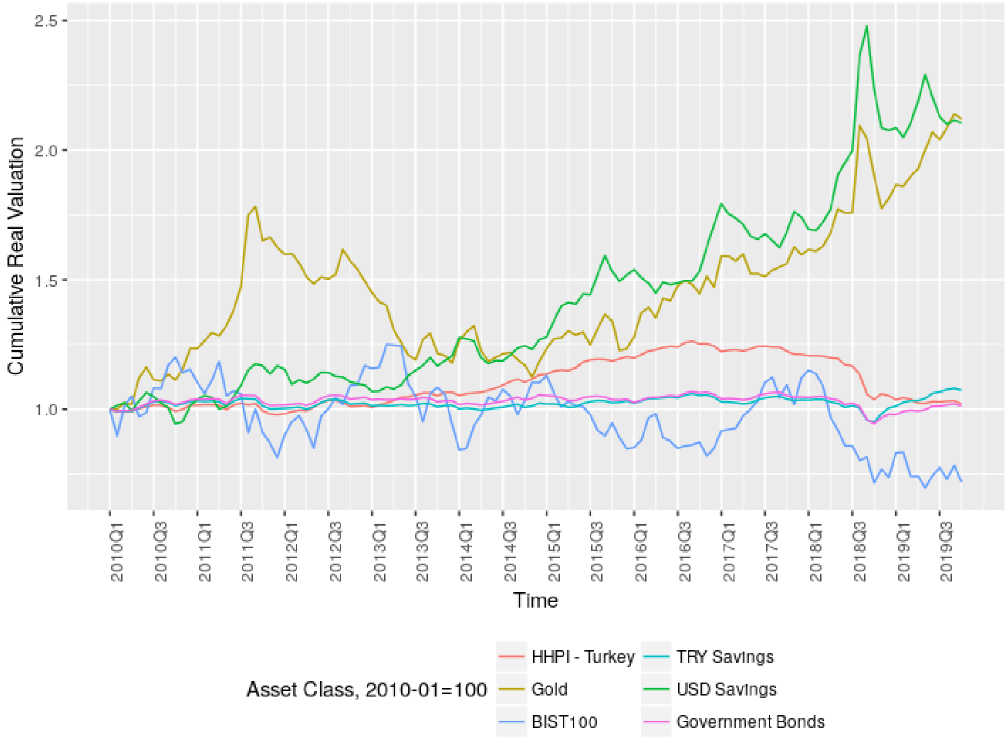

Figure 2. Comparison of housing investment performance versus other asset classes, cumulative, inflation-adjusted, before tax, savings compounded monthly and bonds compounded yearly

Source: Author’s calculation based on CBRT and Thomson Reuters data.

Details on underlying macro trends, such as GDP, CPI, and mortgage rates, as well as significant regulations with an effect on real estate properties are presented in the following sections for a more complete understanding of fundamentals.

1.5 Types of real estate markets and market segmentation

Real estate is a broad term that encapsulates different markets such as housing[5], commercial and land, which differ significantly in their supply and demand dynamics. For a more precise analysis, such grouping can be further decomposed into submarkets such as industrial buildings, retail buildings and office buildings for commercial real estate and single-family houses and condos for housing markets. Segmentation by city, district or neighborhoods is also frequently seen.

There have been previous works on segmentation and grouping of some real estate markets in Turkey. For instance, Baycan and Gulumser (2004) differentiates against four types of gated communities for housing in Istanbul: vertical gated communities, horizontal gated communities, gated apartment blocks and mixed-type gated developments. According to Keskin's (2010) reported expert segmentation, housing in Istanbul is best segmented in 5 submarkets, which are categorized further into groups.

Expert segmentations could be accurate and reliable classifications for social analysis; however, they do not necessarily fit the inherent segmentation in the data and the weights of the submarkets could be overlooked. For instance, although the first submarket in Table 1 contains properties with highest premium, the number of properties there is for most of the subclassifications in this submarket is negligible with respect to common properties and this might show up as an outlier problem in a numeric analysis. Thus, for a broad overlook, a segmentation by listing volume gives a better understanding of the overall market and fits better to a numeric analysis.

[5] The choice of the term "housing" instead of "residential" is to prevent a possible confusion of terms with the Turkish word "rezidans," meaning luxury apartments.

Main segments in the housing real estate market in Turkey, as gathered from the data set of this study and ranked by volume, are shown in Table 2.

Table 1. Keskin's (2010) A Priori Expert Segmentation of Istanbul Housing Market

Housing Submarkets (A priori)	
1st Submarket	1A - Waterside Houses (Yalı) 1B - Horizontal Gated Communities 1C - Vertical Gated Communities 1D - Low storey apartments by the shore, detached houses at the first core of the city
2nd Submarket Apartment Blocks constructed after 1983's neo-liberal economy	2A - Build and Sell Blocks 2B - Sites (Semi Horizontal Gated Community Areas)
3rd Submarket Apartment blocks and detached/attached houses in historical areas.	3A - Apartments blocks in the historical areas 3B - Attached houses in the historical areas
4th Submarket Apartment blocks constructed after 1990's	4A - Build-Sell Apartment blocks 4B - Cooperatives
5th Submarket	5A - Legalized squatter settlements 5B - Squatter settlements 5C - Old summer residential areas

Source: Keskin, 2010..

Similarly using listing volume as a heuristic, and ignoring niches like shopping malls, main submarkets in commercial real estate market in Turkey is as displayed in Table 3.

Table 2. Housing Real Estate Submarkets by Volume

Housing Submarkets by Volume	
1st Submarket Common flats	Attached/detached apartment blocks in cities, social housing
2nd Submarket Squatter settlements	Common flats without proper building permissions, low-profile single houses without proper building permissions or ownership
3rd Submarket High-floor residences	High maintenance, high floor, gated flats in residential projects
4th Submarket Gated houses	Villas, high-profile single family or multiple family houses

Note: Based on Author's calculations.

Submarkets not only differ in their supply and demand dynamics, they are also significantly different in their liquidity. Therefore, they might react differently to macroeconomic shocks. Therefore, while addressing real estate markets, it is important to specify a particular market due to the differences in their pricing dynamics.

Table 3. Commercial Real Estate Submarkets by Volume

Commercial Real Estate Submarkets by Volume	
1st Submarket Retail real estate properties	Stores, showrooms
2nd Submarket Common offices	Older office centers, common flats, small offices
3rd Submarket Industry office buildings	Warehouses, factories, properties in organized industrial areas
4th Submarket Luxury offices	Plazas, newer office centers

Note: Based on Author's calculations.

1.6 Demand, supply, macro trends and regulations in Turkey housing market

A broader market understanding encompasses considerations of underlying trends. For instance, macroeconomic conditions are thought to have an impact on real estate prices through a variety of channels such as mortgage rates, income and inflation. Supply and demand in the market are affected by trends in demographics and credit rates. In this section, relevant figures concerning underlying trends in Turkish Housing Market are presented as a general overview.

As of 2018, Turkey has a population of 82.0 million with a high rate of youth, as the age distribution is shown in the population pyramid in Figure 3. Average household size with a decreasing trend, divorce rates with an increasing trend and marriage rates with a decreasing trend have been supportive of housing demand; while, population increase rate with a slowing trend and birth rates with a decreasing trend might be unsupportive of housing demand in the long run. Figure 4 displays the course of these statistics as provided by Turkstat.

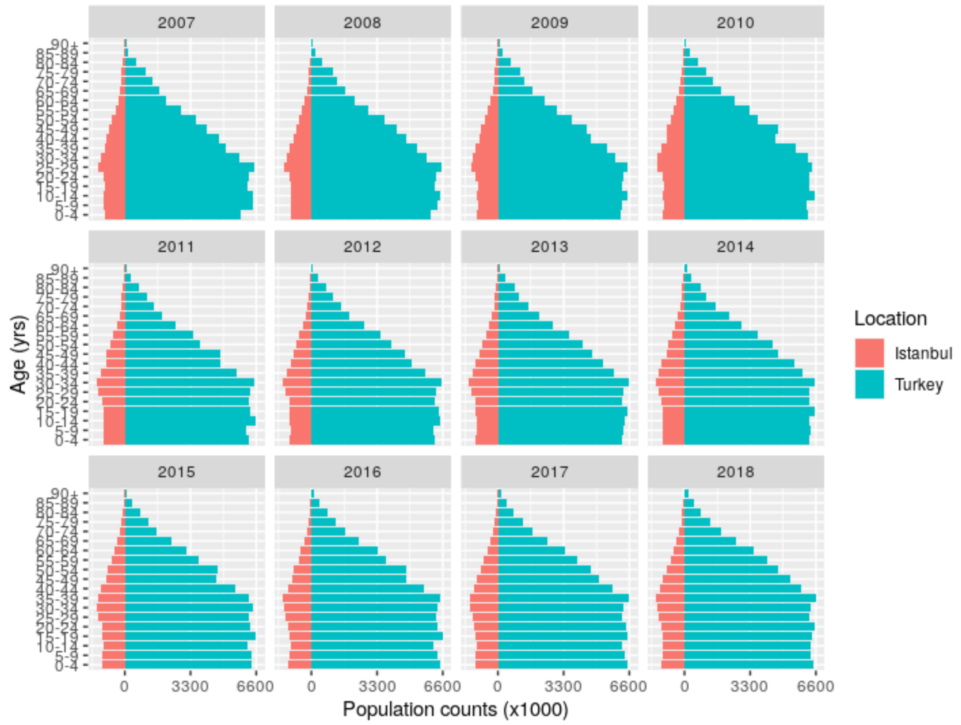

Figure 3 Population pyramid of Turkey and Istanbul by year
Souce: Turkstat data.

Istanbul is the most populous city in Turkey, with a population of 15.1 million[6] living on two continents, with the Bosphorus in between. It has 39 districts

[6] Statistics from Turkstat.

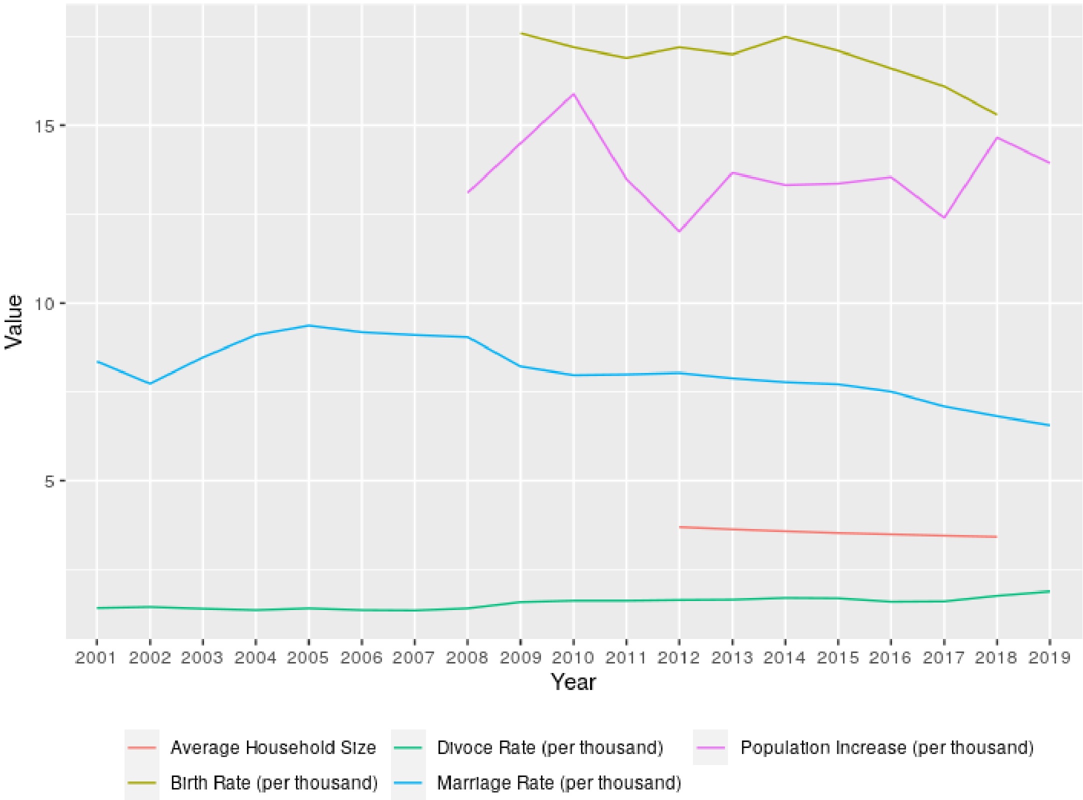

Figure 4 Underlying demographic and social trends in Turkey.
Source: Turkstat data.

and 964 neighborhoods located on a 5461 km2 of land.[7] In 2017, it had received a 1.5% population increase and 238,383 housing units has been sold.[8] It is the economic center of Turkey and traditionally it has received very high rates of migration from other cities; regardless, it also has experienced negative net migration figures in the recent years. Figure 5 shows the population and net migration rates for Istanbul. Macroeconomic factors may also affect housing demand[9] and Figure 6 shows interest rates, consumer price index (CPI) changes, and gross domestic product (GDP) growth, together with real HHPI. One major consideration is that HHPI is available since 2010; regardless, macroeconomic factors historically have

[7] Statistics from Türkiye Mülki İdare Bölümleri Haritası, Harita Genel Müdürlüğü.
[8] Statistics from Turkstat.
[9] See Sari et al. (2007) for a study of the effect of macroeconomic factors in the Turkey housing market.

Figure 6 Population increase and migration rates of Istanbul

Souce: Turkstat data.

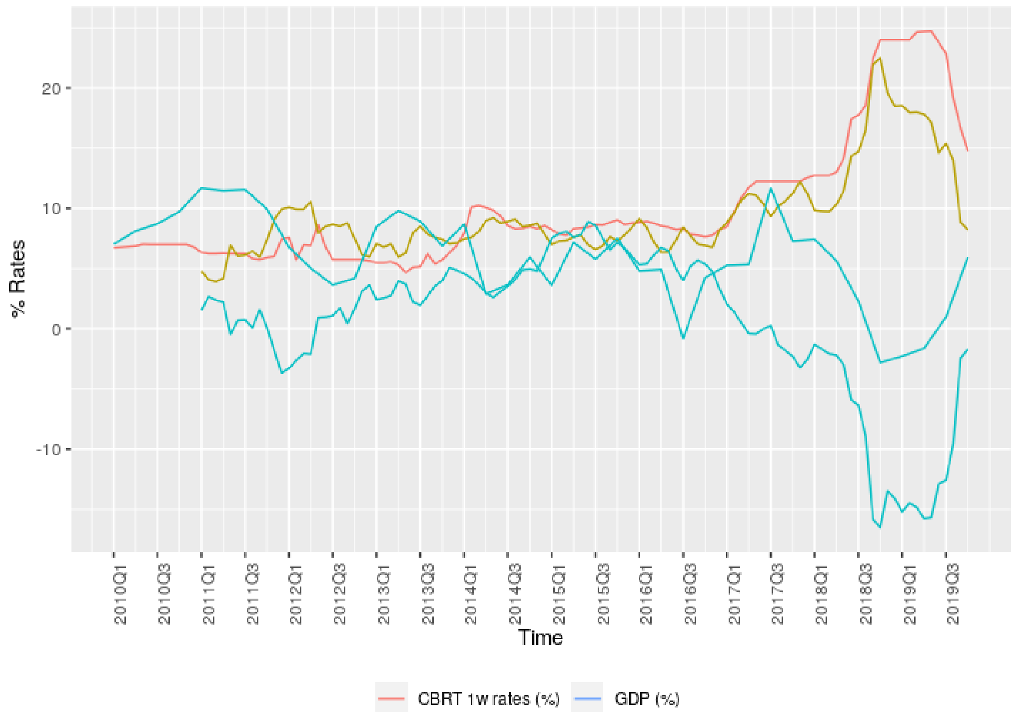

Figure 5 Macroeconomic trends and real HHPI in Turkey

Source: Author's calculations based on CBRT data.

been very volatile for Turkey unlike since 2010. Figure 6 also demonstrates the macroeconomic shock the Turkish economy received in August 2018. Real house prices, being stagnant since late 2016, have slumped with this shock back to initial levels and have not recovered until within the period of analysis.

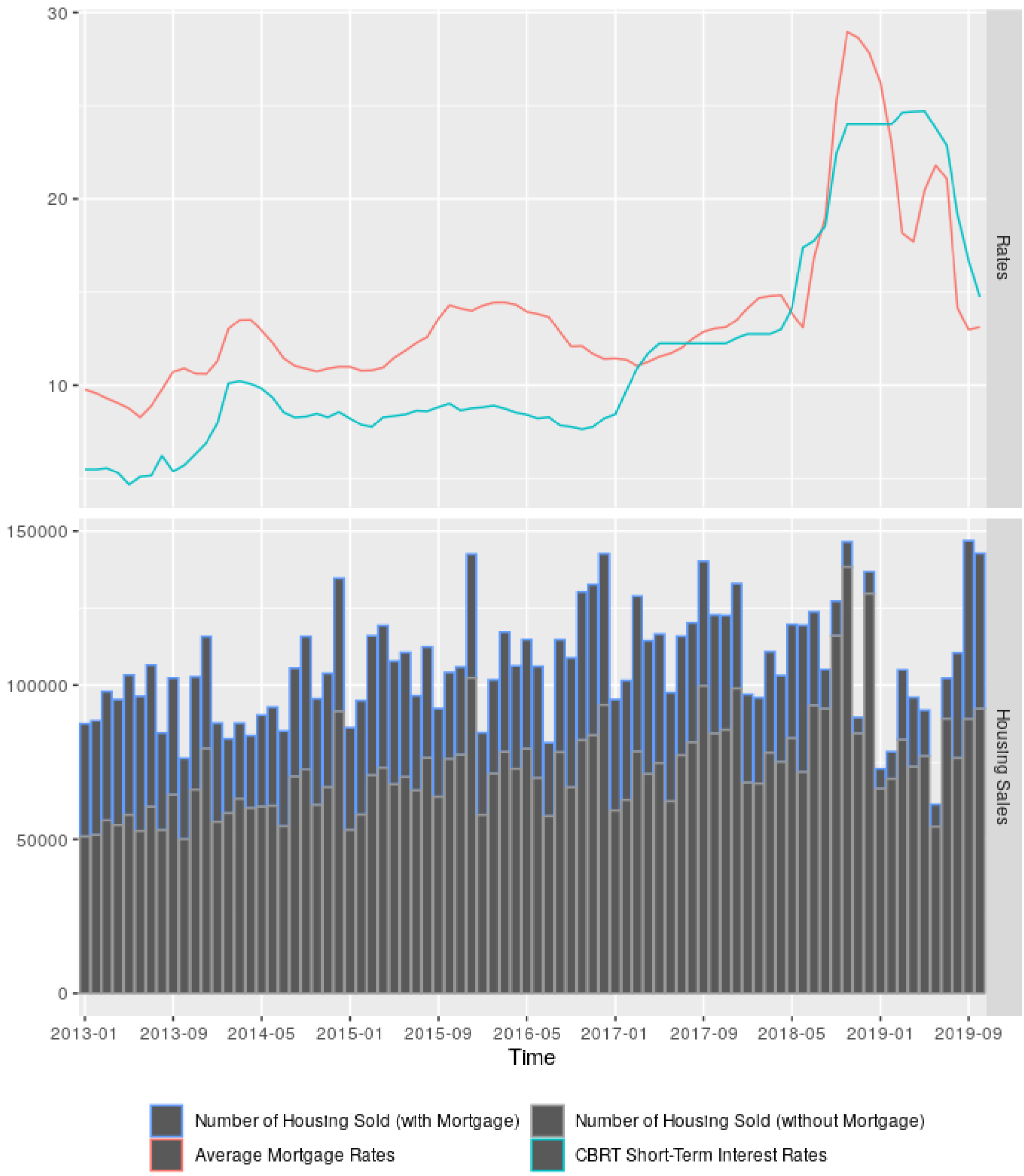

Figure 7 Interest rates, mortgage rates and housing sales in Turkey
Source: CBRT and Turkstat data.

Tied to macroeconomic trends, mortgage rates are a direct factor in housing market demand. Mortgage rates tend to be highly correlated with interest rates as seen in Figure 7. As home ownership rate in Turkey with 67.3%[10] being relatively low and financing could be difficult in an inflationary environment, the mortgage rates directly affect housing sales number, also seen in Figure 7.

It should be noted that even though the mortgage rates significantly affect the demand of housing that is bought with a mortgage, the housing market has shown itself flexible enough to have a margin of suspense as to utilize other sources of credit. One observation is the currency-debt crises of 2018. In the August 2018 outbreak of crises, Turkish Lira has depreciated by 30% and mortgage rates surged to 40%. As Figure 7 shows, this shock caused a dramatic decrease in houses sold with mortgages to drop to near-zero levels and overall housing demand has fallen. Regardless, this was accompanied by the fact that houses bought with other sources have seen a dramatic increase, indicating a flexibility of financing in the market.

As for supply, the housing stock statistics are not publicly available in Turkey; however related construction statistics are provided by Turkstat such as monthly construction permits or use permits, which allow rough estimation of changes in housing stocks. Figure 8 shows the new construction permits each year for housing in Turkey and in Istanbul. In 2017, 15.566 new building permits for housing were issued, becoming the peak construction year in Istanbul. This is a contrast to 10 years prior, 2007, when new housing permits were 2.570. For a more precise interpretation on the housing stock, more detailed figures are needed with a

[10] Turkstat Population and Housing census, 2011. http://www.tuik.gov.tr/PreHaberBultenleri.do?id=15843.

comprehensive study including existing stock, idle housing, lags, and the need for reconstruction of older housing.

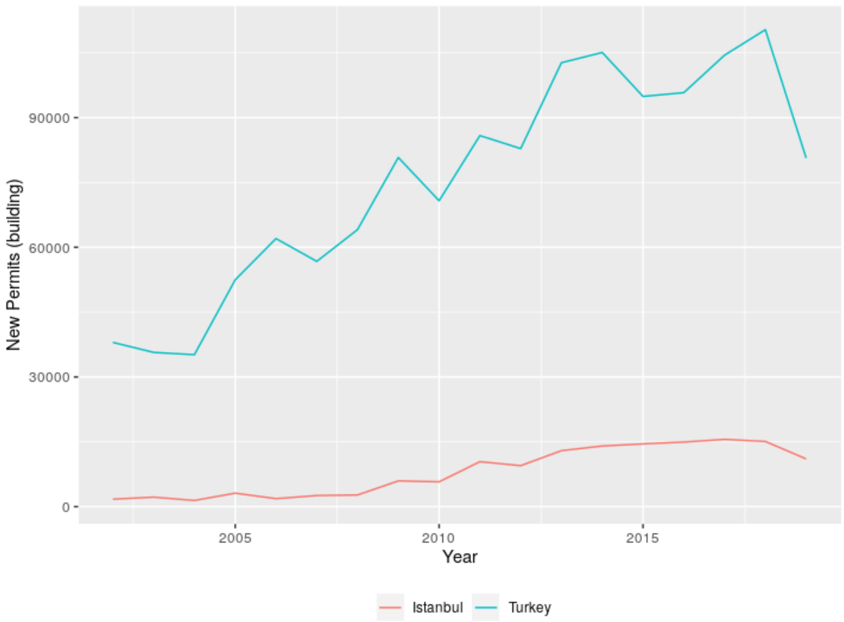

Figure 8 New construction permits for housing in Turkey and in Istanbul.
Source: Author’s calculations based on Turkstat data.

Another aspect to consider in housing supply and demand is the major regulatory interferences and sociological trends over the years, including changes in zonal plans, incentives provided for reconstruction, gentrification efforts and zoning amnesty programs.

Istanbul is expecting a major earthquake in the next few decades and most housing stock built before strengthened regulations are believed to be constructed lacking modern construction approaches. As a consequence, construction permits for most areas are increased for height, to make reconstruction of older buildings profitable by allowing more flats to be reconstructed on the existing land. These regulatory incentives drove a construction boom in Istanbul.

A major regulatory effect on the real estate markets is the city planning. There has been a lack of long-term planning in practice. For instance, there have been 1667 zonal plan changes in Istanbul between 2004 and 2006, while such plans should reflect long-term policies (Turkish Court of Accounts Performance Audit Reports, 2008). Consequently, the housing market can see abrupt price changes in local areas where the plan might focus.

Regulations also have an effect on the market through registration and permits. Istanbul is a city that has seen a rapid increase in population over the last decades and the regulatory bodies were evidently unable to cater for the housing needs of local immigrants. Estimates for buildings with legality issues goes up to include 70% of the buildings in Istanbul, based on expert opinion[11] (Turkish Court of Accounts Performance Audit Reports, 2008). The situation might have been improved over the recent years; regardless it is far from being a minor condition in the real estate markets. This is evident as the zoning amnesty program the government undertook in 2018. With the publishing of the Omnibus Bill numbered 7143 in May 2018, housing (and commercial) property owners with legality issues could apply for registration certificates after paying a fee of 3% (5% for commercial) of the property value as decided by the government.

Gentrification efforts by municipality and government at particular neighborhoods that the government decided to develop also have an effect on the housing supply. Along with a lot of disputes with existing locals, they drove huge booms in supply in local areas where the demand is not yet there. A recent example

[11] Please note that this is not the same as 70% of the population living in buildings with legality issues since multi-floor buildings are more likely to have proper permissions.

is the Fikirtepe neighborhood of Istanbul's Kadikoy district. The neighborhood has been completely changed in a few years from squatter settlements to skyscrapers, not unexpectedly with a high vacancy rate.

One more regulatory effect to real estate markets is the so-called "mega projects" in Istanbul. Istanbul has been an economic center and it attracts a lot of investment, governmental bodies not being an exception. Recent mega projects advertised, planned, conceived, facilitated or realized in Istanbul by governmental bodies include an airport that is among the largest ones in the world, another strait in the city along with a new city settlement around it, new highway system with a new bridge across the Bosphorus, a global financial center and underground tunnels across the Bosphorus. Such large projects, requiring large investments, also affects real estate markets, housing, commercial and land, providing dynamism and shifts in markets.

1.7 Data: sources and description

Lack of reliable data for transaction prices in real estate markets in Turkey is a major bottleneck for research. To my knowledge, there is no public data source available in Turkey that records the real estate transaction prices as it happens. Governmental records are not fully accurate due market practices aiming tax evasion and they are not public, bank records are based on expert valuation rather than transaction data and the data sets of the real estate individual brokerage agencies highly risk non-random sampling.

The approach taken in this thesis is to use listing prices from a major marketplace for real estate listings used by a majority of brokerage agencies. The listing dataset used in this thesis consists of over 300,000 observations of real estate listings across Istanbul as of October 15, 2017, the data provider being HurriyetEmlak.com. HurriyetEmlak.com, which is a well-known one of the two leading online portals and has a large coverage of the market across Turkey. The data set includes all listings in Istanbul that the provider had at the cross-section; therefore, data is considered to be robust to sampling bias and inflationary time effects. The listing dataset is then supported by a primary dataset on information regarding the city with geolocational precision, which enhances the listing dataset by distances to significant city locations.

Each observation in the listing data set consists of an attribute set over 20 variables including features such as asking price, the area in square meters, property types, date added, building age, number of floors in the building, and number of rooms, bedrooms and bathrooms. The data includes multiple real estate markets such as residential, commercial and land markets for sale and for rent. Main housing property types are flats, luxury residences and single houses. Main commercial property types are stores, common offices and luxury offices. The location of an observation is given by the hierarchical administrative areas that the property belongs to in terms of the neighborhood and by the geolocation of the property specifically. The geolocation data allows us to control for spatial autocorrelation, which cannot be fully accounted for at the neighborhood level. Figure 9 shows a mapping of the sale listings for housing in Istanbul.

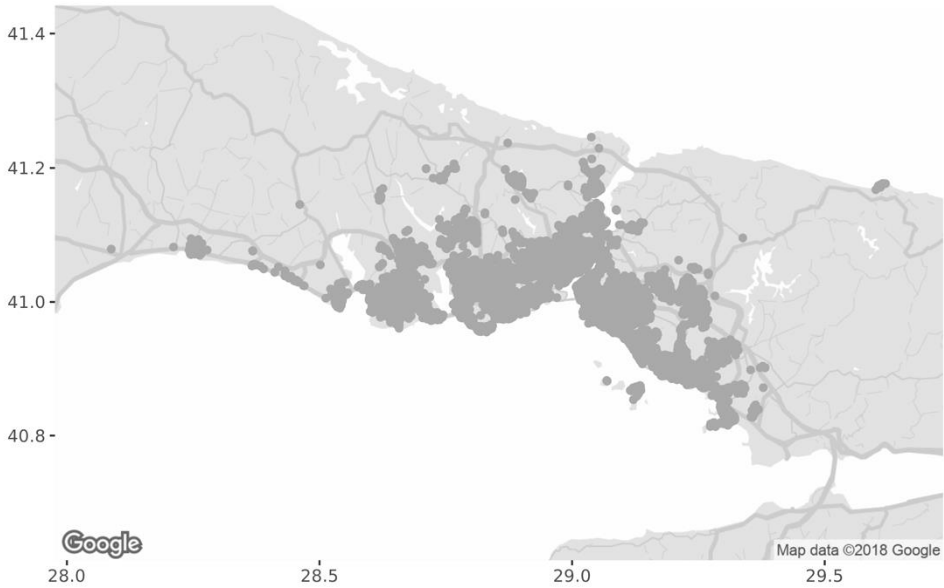

Figure 9 Locations of the sales listings for housing

Maps in Figure 10 colors the housing listings in the dataset by age, price and number of total floors, respectively. A visual examination of the maps gives the insights that the old towns are around Bosphorus, particularly to the south, that prices with a view of Bosphorus around the old towns are much higher, and the newer settlements have higher buildings.

A separate set of data regarding the city, in addition to HurriyetEmlak data is collected to improve the analysis. This set is a primary data set, collected from various official sources that are public on the Internet. This set includes significant places across Istanbul that might have an effect on property prices with their geographical coordinates. Locations of schools, hospitals, police departments, fire departments, shopping malls, universities are collected and distances to these significant places are calculated for all observations.

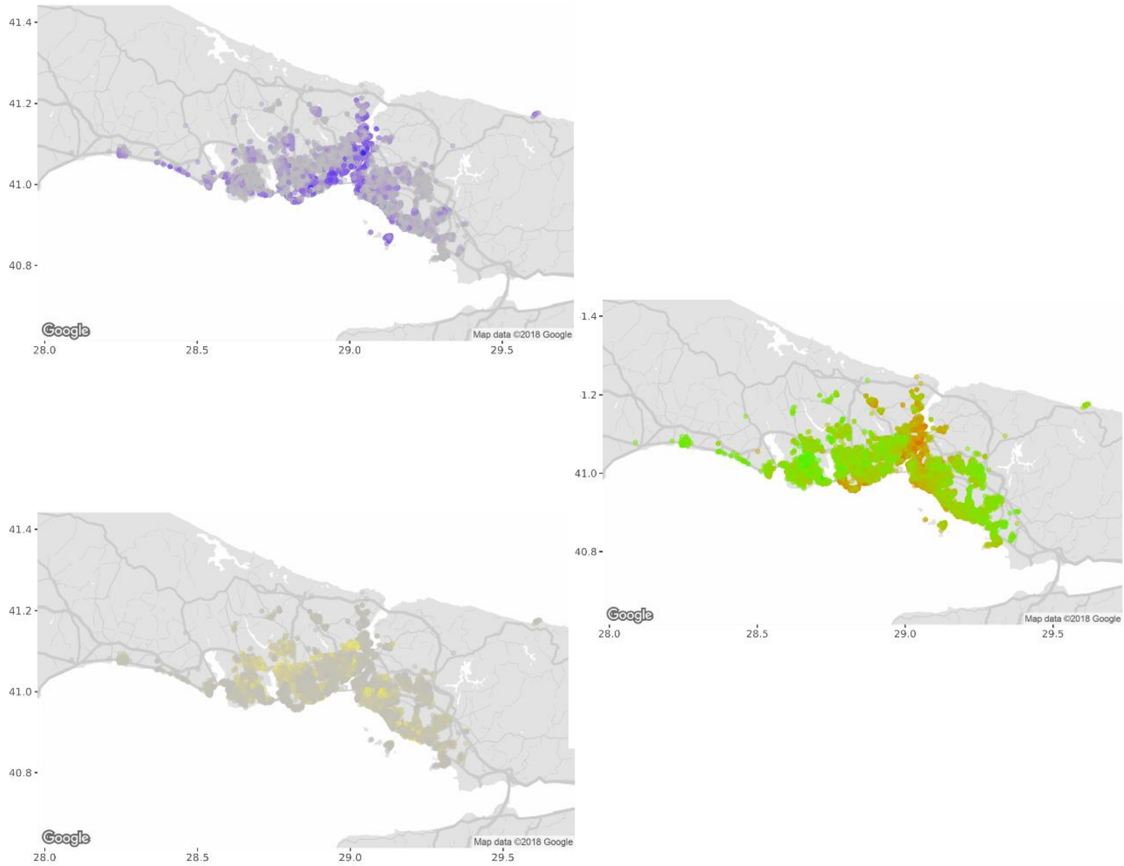

Figure 10 Housing listings for sale in Istanbul, colored by age (top), price (right) and number of total floors (bottom)

The dataset includes listings for both sales and rental and for housing, commercial and land type of real estate markets, making a total of 6 different major segments. A summary statistics of selected numeric variables and categorical variables is provided in Table 4 and Table 5 respectively, for the housing market for sales, after preprocessing of initial data and adding distances to various locations in the city.

Table 5. Descriptives, Istanbul Housing Market, Selected Numeric Variables

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
price	25,219	1,007,670	1,666,121	100,000	340,000	980,000	41,113,600
age	25,219	7.715	11.279	0	0	11	50
area	25,219	129.788	56.101	30	90	150	480
bathrooms	25,219	1.540	0.670	1	1	2	7
livingrooms	25,219	1.057	0.264	0	1	1	4
bedrooms	25,219	2.628	0.975	1	2	3	8
totalFloor	25,219	7.774	6.614	1	4	9	55
dist.2.closest.fire_department	25,219	1,629.150	966.577	6.007	899.813	2,199.070	9,017.453
dist.2.closest.hospital	25,219	920.645	754.680	4.180	455.927	1,147.145	11,067.830
dist.2.closest.shopping_mall	25,219	1,699.333	1,255.944	0.057	864.957	2,300.404	26,485.280
dist.2.closest.police_department	25,219	1,054.478	741.340	6.610	529.276	1,388.698	9,685.606
dist.2.closest.transportation	25,219	1,560.144	2,380.501	6.824	554.241	1,682.442	45,523.780
dist.2.closest.private.preschool	25,219	413.986	475.401	0.847	193.431	515.942	13,834.380
dist.2.closest.private.primschool	25,219	795.807	786.785	4.870	400.891	1,028.156	26,478.820
dist.2.closest.private.highschool	25,219	872.774	818.703	6.414	434.967	1,124.489	26,469.350
dist.2.closest.private.middleschool	25,219	778.555	778.633	5.223	414.033	971.984	26,470.380
dist.2.closest.public.preschool	25,219	2,056.276	1,811.705	9.568	934.576	2,576.942	15,204.980
dist.2.closest.public.primschool	25,219	413.317	300.999	3.033	218.157	522.453	6,730.278
dist.2.closest.public.highschool	25,219	583.254	400.781	2.765	320.106	742.923	8,040.030
dist.2.closest.public.middleschool	25,219	424.949	314.837	4.261	219.509	536.079	5,926.486
dist.2.closest.university	25,219	1,642.899	1,445.822	14.777	755.091	1,977.380	15,328.010

Table 4. Descriptives, Istanbul Housing Market, Selected Categorical Variables

property_type	district	neighborhood	gated	storey_at_clustered	heating_type
Apartman Dairesi:18004	Maltepe :2303	Cumhuriyet : 795	Evet : 1709	middleStorey:6404	Kombi :14810
	Kadky :2182	dealtepe : 601	Hayr:17052	topLevel :3025	Merkezi (Pay ler): 2440
Residence : 757	Bahelievler:1469	Kkyal : 557		firstStorey :2804	Merkezi : 1300
	Beylikdz :1314	Altntepe : 508		groundLevel :2138	Soba : 77
	Kkekmece:1177	Atakent : 490		highStorey :2004	Klima : 57
	Eyp : 924	Adnan Kahveci: 389		halfStorey :1516	Jeotermal Istma : 40
	(Other) :9392	(Other) :15421		(Other) : 870	(Other) : 37

1.8 Methods and tools

This thesis adopts social science research methodologies with an emphasis on data. The data gathered across various sources are scrutinized from multiple points of view in order to find evidence for or against hypotheses under consideration. Statistical hypothesis testing and econometric models are used with a frequentist approach in data analysis and a distributed computation power that is outside the capabilities of a single home computer are used from time to time.

While the time-series econometric models are the most common models in finance as they are suited to work with time-dependent financial data and returns; the nature of the data of this thesis demands a different set of econometric models. Particularly, cross-sectional spatial econometric methods are used, for which the next chapter is dedicated. Further details specific to research questions under consideration are provided in relevant sections of later chapters as well.

Open-source R software and its GUI R-Studio is used for the statistical analyses performed in this thesis. It should be noted that this thesis rests upon a vast amount of open source development effort. In addition to the core R libraries, I made use of many libraries for whether in statistical analysis or as helpers such as in data manipulation or visualizing outputs, including: "readr", "lubridate", "geosphere", "stargazer", "gstat", "spdep", "spatialEco", "ggmap", "ggplot2", "gridExtra", "tictoc", "googleVis", "ggmap", "RgoogleMaps", "corrplot", "automap", "multiway", "Hmisc", "sphet", "caTools", "scales", "stringr".

1.9 Dissertation summary and policy implications

Dynamics of pricing in real estate markets is studied throughout this dissertation using micro data with geolocational precision. Spatial models and tools for pricing are explained in detail. An analysis of space in housing markets is provided and the theory of the causes of spatial dependence is constructed and tested. Additionally, pricing behavior of people is examined through an analysis of pricing distributions in multiple real estate markets.

Spatial dependence is found not only in cross-sectional real estate prices, but also in structural and apartmental variables of the properties and in pricing model residuals, While the spatial dependence in residuals is caused by omitted variables or model misspecification; the process by which it happens in housing market is further analyzed. Results of this study found empirical support for the hypotheses that the construction process, shared social services and high-rise luxury residences are determinants of spatial autocorrelation.

Behavioral effects of ‘even pricing’ and ‘precision avoidance’ are detected in price series. It is documented that a course pricing set clustered around the numbers 0 and 5 is used in real estate and it is statistically shown that this phenomena constitutes an anomaly in distributions and pricing models at an economically significant level. In further analysis of the determinants, it is found that solely rational explanations such as price levels, information frictions or taxes are insufficient to explain the observed phenomena. Empirical findings of this study are in line with cognitive ease and easier negotiation process hypotheses and furthermore a cultural difference in pricing behavior is found on charm pricing.

This dissertation is relevant to contemporary literature in multiple folds. It shows where and how standard hedonic pricing models fail. Not only it demonstrates how to account for space in analysis, it also shows in detail what happens when ignored and what proxies can be used to account for space. Moreover, it extends behavioral approaches in finance to real estate markets and provides a groundwork for possible regulatory actions against behavioral inefficiencies in real estate markets. Additionally, it lays out where theories regarding real estate markets in developed markets apply also in developing markets and where they do not align.

The research done has some direct policy implications. First, the economic environment and the dynamics of economics in emerging markets could be different than developed markets and therefore automatic adopting of policies is not appropriate. Market differences should be considered in policy making. For instance, any policy decision or time-series analysis in real estate markets in Turkey should consider the high number of zonal changes to be effective. Second, not only geography affects real estate prices, it also can be an issue for our pricing models. The use of spatial econometric techniques can help forming more accurate models and allows identifying possible problems. For instance, the cultural differences between gated high-rise buildings and its surroundings is reflected in and identified through spatial autocorrelation of pricing models and policy decisions should take the differences between high-rise buildings and its surroundings into consideration. Third, while financial markets are the natural home for financial research, financial research is not limited to financial markets and extends to real markets as well. This includes behavioral biases, therefore policies countering behavioral inefficiencies are called for. Findings of this study suggest that financial regulations that are based on behavioral biases could have parallels in the regulations on real estate markets, therefore more research is needed to make the regulations in multiple markets aligned, founded and simpler.

Examining micro data in-depth in multiple studies unravel some indirect policy implications as well. First, while ideal policy making should rely on facts and quality research, real estate markets in Turkey seem to be suffering from a lack of data. The lack of quality data presents a bottleneck for quality research, thus efficient policy making. Particularly, accuracy and coverage of transactional data for an extended period is currently missing in the Turkish real estate market, solving which would enable better policy making.

Another implication regards the variation in prices. The prices vary greatly across the city, suggesting a huge wealth disparity. If we treat house prices as a wealth proxy and grossly assume that each house represents a single household's wealth, the resulting Gini coefficient is 0.536, a very high inequality number compared to Turkey's Gini coefficient for income inequality (0.419). As we know that the homeownership rate is less than 1 (around 67%) and houses do not represent individual households but rather are concentrated among fewer owners, the wealth disparity in the public seems to be much worse, calling for general policies against wealth inequality.

A final indirect implication regards regulation imbalance. The real estate market in Turkey seems to be regulated extensively, while a majority of housing properties in Istanbul suffer from permit issues. Moreover, the number of zone planning changes seems strikingly high, indicating a lack of long-term planning at the minimum. While there is certainly not a lack of regulatory effect, on the other hand, regulations regarding behavioral market inefficiencies are not in place. Such observations point out to an evident policy inefficiencies in real estate markets, which calls for, perhaps fewer but more effective regulations.

CHAPTER 2

EXPLORING SPATIAL MODELS FOR THE CROSS-SECTION

2.1 Introduction

There are various ways to incorporate space into analysis in a regression. Adding binary variables indicating administrative units to account for locational fixed effects is a common way. Calculating distances to significant places, such as central business districts or shopping malls is also frequently applied. Moreover, space is inherently incorporated in many independent variables that are seemingly non-locational variables. In contrast, another set of ways of incorporating space is to use spatial econometric models. Methods listed in this section are some of the econometric techniques to model spatial dependence in regressions. This chapter presents spatial econometric models that could be used with cross-sectional data.

The aim of this chapter is to lay out some spatial econometric tools and models available for the cross-section as they are not commonly used in financial literature. Moreover, as both available models and their terminology is quite rich and variable (and sometimes overlapping) in literature of spatial econometrics, this chapter will be used as a reference on spatial models in the rest of this dissertation, facilitating a more precise discussion on spatial dependence.

Spatial models have been named and abbreviated in multiple ways in the literature and the naming could sometimes refer to slightly or significantly different models.[12] The naming convention presented here for the names of the spatial models

[12] The terminology could be confusing in the literature. In addition to the way a SAR model is presented here as a pure autoregressive function, it is frequently used to refer to SLM such as by LeSage and Pace. Bivand, Pebesma,& Gómez-Rubio uses SAR abbreviation as an abbreviation for simultaneous autoregressive models as opposed to spatial autoregressive models. Anselin uses SAR

and their abbreviations are selected carefully from the textbooks of Anselin (1988), Cressie (1993), LeSage & Pace (2009), Bivand, Pebesma, & Gómez-Rubio (2013), and Arbia (2014) in order to increase clarity.

An important distinction in terminology is related to explaining spatial relations between variables. Anselin (1988) separates spatial effects into two: spatial dependence (spatial autocorrelation) and spatial heterogeneity. He defines spatial dependence as "the lack of independence which is often present among observations in cross-sectional data sets" and spatial heterogeneity as "structural instability over space" such as "heteroskedasticity, random coefficient variation and switching regressions". This study follows the terminology definitions in his book (1988) and differentiates between spatial dependence and spatial heterogeneity while does not differentiate between spatial dependence and spatial autocorrelation.

Spatial models are explained in detail next and a discussion is given in the conclusion afterwards.

2.2 Spatial lag and spatial autoregression

Consider a variable of concern, v, and a spatial definition of weight matrix form, W.

The term *spatial lag* refers to the matrix product Wv.

terminology in referring to SEM model with autoregressive errors. Arbia uses SLM to refer to SDM. SDM is sometimes referred to as spatial common factor models. SARMA models are sometimes referred to as SARAR or SAC models, with or without nuances. Examples of different uses of terminology are plenty.

The definition of a *spatial autoregression* would naturally follow to explain a variable by its spatial lag as in:

$$v = Wv + \epsilon \tag{1}$$

Regardless, I will not use the term spatial autoregressive model (SAR) to refer to this function due to its loaded use in the literature to refer to various models; but, instead use *spatial autoregression*. Particularly note that the term SAR can be used in literature to refer to (a) a pure spatial autoregression model as explained in this section, (b) a SLM model explained in the next section which also could include additional exogenous variables, (c) a SEM model where the disturbances are modeled in an autoregressive manner and (d) simultaneous autoregressive models.

2.3 Spatial lag model (SLM)

Spatial lag model refers to models where a spatial lag of the dependent variable is used as an explanatory variable:

$$y = \alpha + \lambda W y + \epsilon, |\lambda| < 1 \tag{2}$$

$where$

$y \text{ is the dependent variable vector,}$

$\alpha \text{ is the intercept,}$

$\lambda \text{ is a scalar autocorrelation parameter,}$

$W \text{ is the spatial weight matrix,}$

$and\ \epsilon\ is\ the\ random\ error\ term\ vector.$

While the model could be a pure spatial autoregression model as above, exogenous independent variables can also be added. A more complete SLM model refers to:

$$y = \alpha + \lambda Wy + X\beta + \epsilon, |\lambda| < 1 \tag{3}$$

$where$

$y\ is\ the\ dependent\ variable\ vector,$

$\alpha\ is\ the\ intercept,$

$\lambda\ is\ a\ scalar\ autocorrelation\ parameter,$

$W\ is\ the\ spatial\ weight\ matrix,$

$X\ is\ a\ set\ of\ independent\ variable\ vectors,$

$\beta\ is\ a\ set\ of\ coefficients,$

$and\ \epsilon\ is\ the\ random\ error\ term\ vector$

SLM can be estimated using maximum likelihood estimation or 2SLS methods.

2.4 Spatial error model - weight matrix specification (SEM)

While SAR and SLM model the spatial autoregressive relation within the dependent variable, *spatial error model (SEM)* models the spatial autoregressive relation within the residuals. Following Dubin's presentation (1998), consider a linear model:

$$y = \alpha + X\beta + u \, , \tag{4}$$

$where$

$y\ is\ the\ dependent\ variable\ vector,$

$\alpha\ is\ the\ intercept,$

$X\ is\ a\ set\ of\ independent\ variable\ vectors,$

$\beta\ is\ a\ set\ of\ coefficients,$

$and\ u\ is\ the\ residual\ vector$

The residuals are spatially modeled with respect to a spatial weights matrix:

$$u = \lambda W u + \epsilon \, , \tag{5}$$

$where$

$W\ is\ the\ spatial\ weight\ matrix,$

$\lambda\ is\ a\ scalar\ autocorrelation\ parameter,$

$and\ \epsilon\ is\ the\ random\ error\ term$

Solving above equation for u:

$$u \ = (I \ - \lambda W)^{-1} \epsilon \tag{6}$$

Substituting the residual vector back to the original equation, the spatial error model can be represented in reduced form:

$$y = \alpha + X\beta + (I \ - \lambda W)^{-1} \epsilon \, , \tag{7}$$

$where$

$y \text{ is the dependent variable vector},$

$\alpha \text{ is the intercept},$

$X \text{ is a set of independent variable vectors},$

$\beta \text{ is a set of coefficients},$

$I \text{ is the identity matrix},$

$\lambda \text{ is a scalar autocorrelation parameter},$

$W \text{ is the spatial weight matrix},$

$\text{and } \epsilon \text{ is the random error term}$

Variance-covariance matrix is given by:

$$V = E[uu'] = \sigma^2 (I - \lambda W)^{-1} (I - \lambda W')^{-1} \qquad (8)$$

The weight matrix model is a priori modelling of the weight matrix W by the researcher to represent spatial effect of an instance on another. Spatial representation in W has been done by various ways such as binary neighborhood relations, kth-nearest-neighbor relations and distance-based relations.

SEM models can be estimated using maximum likelihood estimation or feasible GLS estimation (Arbia, 2014).

2.5 Spatial error model - direct covariance specification

As an alternative to specifying weights and deriving the covariance matrix in a SEM, a functional form of covariance structure can be assumed (Dubin, 1998). This direct covariance specification is typically a function of distances. An example would be the Gaussian form as Equation 9:

$$K_{ij} = b_1 exp(-\frac{D^2{}_{ij}}{b_2}) \text{ ,} \tag{9}$$

where

K_{ij} *is the covariance between observations* i *and* j,

b_1 *and* b_2 *are constants,*

and D_{ij} *is the distance between observations* i *and* j

2.6 Spatial error model - spatial moving average (SMA)

The spatial dependence in the error terms could also take the form of a moving average process, resulting in a *spatial moving average model (SMA)*:

$$y = \alpha + X\beta + u \text{ ,} \tag{10}$$

where

y *is the dependent variable vector,*

α *is the intercept,*

X *is a set of independent variable vectors,*

β *is a set of coefficients,*

and u *is the residual vector*

The residuals are spatially modeled with respect to a spatial weights matrix in the form of a moving average process:

$$u = \lambda W\epsilon + \epsilon \,, \tag{11}$$

where

λ *is a scalar autocorrelation parameter,*

W *is the spatial weight matrix,*

and ϵ *is the random error term*

2.7 Spatial lag of X model (SLX)

While SLM and SEM models take the spatial dependence in the dependent variable and residuals into account, respectively, *spatial lag of X model (SLX)* incorporates the spatial lags of the independent variables. The SLX model is given as (Vega & Elhorst, 2015):

$$y = \alpha + X\beta + WX\theta + \epsilon, \tag{12}$$

where

y *is the dependent variable vector,*

α *is the intercept,*

$X \text{ is a set of independent variable vectors},$

$\beta \text{ and } \theta \text{ are sets of coefficients},$

$\text{and } \epsilon \text{ is the random error term}$

SLX model is also called spatial cross-regressive model in the literature. (Florax & Folmer, 1992)

2.8 Spatial Durbin model (SDM)

The *spatial Durbin model (SDM)* is first given by Anselin (1988) as:

$$y = \alpha + \lambda W y + X\beta - \lambda W X \beta + \epsilon \tag{13}$$

where

$y \text{ is the dependent variable vector},$

$\alpha \text{ is the intercept},$

$\lambda \text{ is a scalar autocorrelation parameter},$

$W \text{ is the spatial weight matrix},$

$X \text{ is a set of independent variable vectors},$

$\beta \text{ is a set of coefficients},$

$\text{and } \epsilon \text{ is the random error term}$

SDM incorporates spatial lags of both the independent and the dependent variables. A spatial Durbin model can be reached treating a SEM model as a lag model. Recall that the SEM model is given by the reduced form:

$$y = X\beta + (I - \lambda W)^{-1}\epsilon\,, \tag{14}$$

$where$

$y\ is\ the\ dependent\ variable\ vector,$

$X\ is\ a\ set\ of\ independent\ variable\ vectors,$

$\beta\ is\ a\ set\ of\ coefficients,$

$I\ is\ the\ identity\ matrix,$

$\lambda\ is\ a\ scalar\ autocorrelation\ parameter,$

$W\ is\ the\ spatial\ weight\ matrix,$

$and\ \epsilon\ is\ the\ random\ error\ term$

Rearranging the terms in the reduced form, we get:

$$(I - \lambda W)\, y = (I - \lambda W)\ \ X\beta + \epsilon$$

$$y - \lambda W y = X\beta - \lambda W X\beta \quad + \epsilon$$

$$y = \lambda W y + X\beta - \lambda W X\beta \quad + \epsilon \tag{15}$$

The SDM approach, in contrast, can account for the correlation between the covariates and the error term. In other words, while SLM and SEM models do not account for the spatial correlation between the independent variables and the disturbance, spatial Durbin model incorporates spatial lags of the independent variables into the regression as well. To see this, assume there is a linear correlation between the errors and the covariates in SEM, modeled by the correlation coefficient θ.

$$y = X\phi + (I - \lambda W)^{-1}u$$

$$u = \theta X + \epsilon \tag{16}$$

When we substitute the correlated error vector back to the original regression and rearrange terms, we once again get the SDM model.

$$y = X\phi + (I - \lambda W)^{-1}(\theta X + \epsilon)$$

$$(I - \lambda W)y = (I - \lambda W)X\phi + \theta X + \epsilon$$

$$y = \lambda Wy + X\phi - \lambda WX\phi + \theta X + \epsilon$$

$$y = \lambda Wy + X(\phi + \theta) + WX(-\lambda\phi) + \epsilon$$

$$y = \lambda Wy + X\beta + WX\gamma + \epsilon \tag{17}$$

This form is called the unconstrained spatial Durbin model (LeSage & Pace, 2009) and it is a generalized version of SEM, where SEM is a special case of SDM when $\gamma = -\lambda\phi$. Whether an estimated SDM model is in fact points to a SEM model can be formally tested by a common factor test (Anselin, 1988), where the null hypothesis is that there is a common factor, that is, $H_0: \gamma = -\lambda\beta$. If the test is rejected, then the evidence favors SDM rather than SEM.

Spatial Durbin model can be estimated with maximum likelihood or 2SLS estimation methods (Arbia, 2014).

2.9 Spatial Durbin error model (SDEM)

Another way of modelling spatial effects is by adding spatial lags of the covariates to a SEM, which is called a *spatial Durbin error model (SDEM)* (LeSage & Pace, 2009). SDEM is given by the equations:

$$y = \alpha + X\beta + WX\gamma + u$$

$$u = \lambda Wu + \epsilon , \tag{18}$$

where

y *is the dependent variable vector,*

α *is the intercept,*

X *is a set of independent variable vectors,*

β *is a set of coefficients,*

W *is the spatial weight matrix,*

γ *and* λ *are scalar autocorrelation parameters,*

u *is the residual vector*

and ϵ *is the random error term*

In reduced form, SDEM becomes:

$$y = \alpha + X\beta \ + WX\gamma + (I - \lambda W)^{-1}\epsilon \tag{19}$$

The difference between SDM and SDEM is that SDM includes a spatial lag of the dependent variable, while SDEM replaces that with spatially dependent errors.

2.10 Spatial autoregressive with additional autoregressive error (SARAR)

Another set of models account for spatial dependence both in covariates and in the error terms. Such models are called *spatial autoregressive with additional autoregressive error (SARAR)* models and defined by the following equations:

$$y = \alpha + \lambda W_1 y + X\beta + u$$

$$u = \rho W_2 u + \epsilon \tag{20}$$

where

y *is the dependent variable vector,*

α *is the intercept,*

X *is a set of independent variable vectors,*

β *is a set of coefficients,*

W_1 *and* W_2 *are spatial weight matrices,*

ρ *and* λ *are scalar autocorrelation parameters,*

u *is the residual vector*

and ϵ *is the random error term*

This model is a generalized version of SLM and SEM models: if ρ is zero, SARAR reduces to SLM model and if λ is zero, it reduces to the SEM model. This model is also referred to as SAC (spatial autoregressive combined), SARSAR, or general spatial model in the literature.

The spatial weight matrices may or may not be the same, without any change on parameter interpretations; however, it is possible to run identification problems to differentiate against λ and ρ, if they are the same.

SARAR model can be estimated by maximum likelihood, GS2SLS or Lee's IV estimation (BFGS2SLS) methods, albeit with possible efficiency problems (Arbia, 2014).

2.11 Spatial autoregressive moving average model (SARMA)

The spatial error structure in a generalized spatial model could be in the form of moving averages as well. A SARMA model is given by the following equations:

$$y = \lambda W_1 y + X\beta + u$$

$$u = \rho W_2 \epsilon + \epsilon \quad (21)$$

where

y is the dependent variable vector,

α is the intercept,

X is a set of independent variable vectors,

β *is a set of coefficients,*

W_1 *and* W_2 *are spatial weight matrices,*

ρ *and* λ *are scalar autocorrelation parameters,*

u *is the residual vector*

and ϵ *is the random error term*

In reduced form, the model becomes:

$$y = \lambda W_1 y + X\beta + (1 + \rho W_2)\ \epsilon \tag{22}$$

2.12 Generalized nesting spatial model (GNS)

An even higher order set of models can be devised to include all spatial interactions. Vega and Elhorst (2015) label this model as *generalized nesting spatial model (GNS).*

$$y = \alpha + \lambda W_1 y + X\beta + W_1 X\theta + u$$

$$u = \rho W_2 u + \epsilon \tag{23}$$

where

y *is the dependent variable vector,*

α *is the intercept,*

X *is a set of independent variable vectors,*

β *and* θ *are sets of coefficients,*

W_1 and W_2 are spatial weight matrices,

ρ and λ are scalar autocorrelation parameters,

u is the residual vector

and ϵ is the random error term

2.13 SLM - matrix exponential spatial specification (MESS)

Matrix calculations are computationally expensive procedures, as for instance the time complexity of an eigenvalue calculation is $O(n^3)$ (Anderson et al., 1999) and the space complexity of keeping a spatial weights matrix is $O(n^2)$. While the complexities do not pose a problem for small samples, for larger samples computational concerns rise.

LeSage and Pace (2007) proposes a *matrix exponential spatial specification (MESS)* to address the computational issues as an alternative to SLM. The SLM model is transformed with a matrix exponential spatial process, resulting in an exponential decay in the effect of the higher-order neighbors rather than a geometric decay.

The MESS model is given by:

$$Sy = X\beta + \epsilon \tag{24}$$

where

S is a real positive definite matrix,

y *is the dependent variable vector,*

X *is a set of independent variable vectors,*

β *is a set of coefficients,*

ϵ *is the random error term vector*

Specifying S as follows yields a model alternative to SLM:

$$S = e^{\alpha W} = \sum_0^\infty \quad \frac{a^i W^i}{i!} \tag{25}$$

where

α *is a scalar real parameter*

W^i *are spatial weight matrices, representing neighboring relations of order* i

Although the MESS estimation is introduced as an alternative to SLM, it is not particular to it and could be extended to SEM models (LeSage & Pace, 2009).

2.14 Geographically weighted regression (GWR)

Brunsdon, Fotheringham & Charlton (1996) proposes a different technique for incorporating a space definition to regressions, called *geographically weighted regression (GWR)*. This model explicitly allows parameter estimates to vary over space and thus differs from the methods above in not modelling residuals.

Bitter et al (2007) provides a good representation of GWR. Consider a standard hedonic regression, in a representation where the subscripts are explicitly denoted:

$$y_i = \alpha_i + \sum_k \quad \beta_k x_{ik} + \epsilon_i \ , \tag{26}$$

$where$

$\alpha \ is \ the \ intercept,$

$\beta_k \ is \ the \ coefficient \ of \ k^{th} \ independent \ variable$

$x_{ik} \ is \ the \ i^{th} \ observation \ of \ k^{th} \ independent \ variable$

$and \ \epsilon \ is \ the \ random \ error \ term$

GWR differs from standard hedonic regression by making coefficients local estimates instead of a global estimate.

$$y_i = \alpha_i + \sum_{k=1,m} \quad \beta_{ik} x_{ik} + \epsilon_i \ , \tag{27}$$

$where$

$\alpha \ is \ the \ intercept,$

$\beta_{ik} \ is \ the \ coefficient \ of \ the \ k^{th} \ independent \ variable \ at \ location \ i$

$x_{ik} \ is \ the \ i^{th} \ observation \ of \ k^{th} \ independent \ variable$

$and \ \epsilon \ is \ the \ random \ error \ term$

The variation in a parameter's coefficient estimates across observations comes from a weighted least squares approach such that different emphasis is put on different observations during estimation of a parameter at different locations. Thus, the estimator of GWR in matrix notation is:

$$\widehat{\beta_i} = (X' \ \ G_i \ X)^{-1} \ X' \ \ G_i \ y \tag{28}$$

$where$

X $is\ a\ set\ of\ independent\ variable\ vectors,$

G_i $is\ a\ spatial\ weighting\ matrix$

y $is\ the\ dependent\ variable\ vector$

Note that the spatial weighting matrix in GWR represents the spatial definition through which the weights of observations around a locality in a weighted regression are determined. This is in contrast to spatial weights matrices in above models, where it represents the spatial definition among observations through which a spatial process occurs and spatial interaction is observed in the form of error structure or spillovers. Both matrices, on the other hand, serve the same purpose of favoring nearby observations in a prediction setting. One possible specification for the spatial weighting matrix in GWR could again be a Gaussian form:

$$G_{ij} = exp(-d_{ij}/h)^2 \qquad (29)$$

$where$

d_{ij} $is\ the\ distance\ between\ observations\ i\ and\ j$

h $is\ the\ bandwidth\ parameter$

The bandwidth parameter is a tool to control the window of the sliding regressions over the space. The larger the bandwidth gets, the larger the range of the nearby observations included is and the closer the the GWR model to an OLS model. The model gets more local as the bandwidth gets smaller; however, the locality is bounded by the fact that the number of observations should be high enough for each regression to provide enough degrees of freedom for estimation. While a priori specification of the spatial weighting schema, the kernel, and the bandwidth is required in GWR, there are adaptive procedures through which a bandwidth is selected in an automatic manner.

GWR fits as many weighted regressions as the observations and it is a tool for working with spatial heterogeneity, where regression coefficients are expected to largely vary across a surface.

2.15 Further extending the models

While models described above constitute a good selection, it is not a comprehensive list of all available models for the cross-section. Making a comprehensive list is a diligent work not attempted here; however, there are particularly three sets of extensions I would like to mention for completeness. The combination of these extensions with presented ones constitutes the vast majority of contemporary cross-sectional techniques.

Importantly, the models listed above rely on the assumption that the random error terms have constant variance. Spatial analysis, on the other hand, might result in heteroscedastic errors, especially while working in large areas that could contain various local spatial processes. To overcome the problem of variable variance of errors, spatial HAC estimators that are robust to heteroscedasticity are proposed in the literature (Kelejian & Prucha, 2007), opening up a new set of spatial models, that is *heteroscedastic spatial models*.

Another set of spatial models comes with limited dependent variables. *Discrete choice models*, working with truncated or censored data are also possible in spatial settings (LeSage & Pace, 2009).

Finally, a Bayesian approach can be taken with spatial models (Lesage, 1997; LeSage, 2000; LeSage & Parent, 2007). Markov Chain Monte Carlo simulations are used in numerical estimations of models.

2.16 Discussion & conclusion

This study presented a list of available models and techniques to account for sorts of spatial dependence for cross-sectional data. The tools given in this section should facilitate a better discussion in next chapters in terms of pinpointing terms and being a reference.

The spatial models presented here can be classified in multiple folds. A major classification is about the kind of spatial relation the model is focused on. Anselin (1988) defines two types of spatial relation: spatial heterogeneity and spatial dependence. Spatial heterogeneity is about instability over space and it is related to what could be referred to as non-stationarity over space. The main tool available to approach spatial heterogeneity is the GWR method, where the method allows the coefficient estimates to vary across space. All other models mentioned in this study deal with spatial dependence, also known as spatial autocorrelation, which is defined as the lack of dependence among observations in a spatial (cross-sectional) setting.

Some spatial models discussed are sometimes referred to as higher order models based on the amount of spatial components and interactions they include. For instance, SEM or SLM are simpler spatial models as they only include one spatial component, that is spatial error or spatial lag (respectively). In contrast SDE or SDEM are higher order models as they can be considered as a generalization of a SEM model. A GNS model is even a higher order model as SDE and SDEM can be considered specific instances of a GNS model. The relationships between many of the models above is graphically displayed in Figure 11, where the higher and lower order models and their parametric relation in perspective is given following the Figure 1 of Vega & Elhorst (2015).

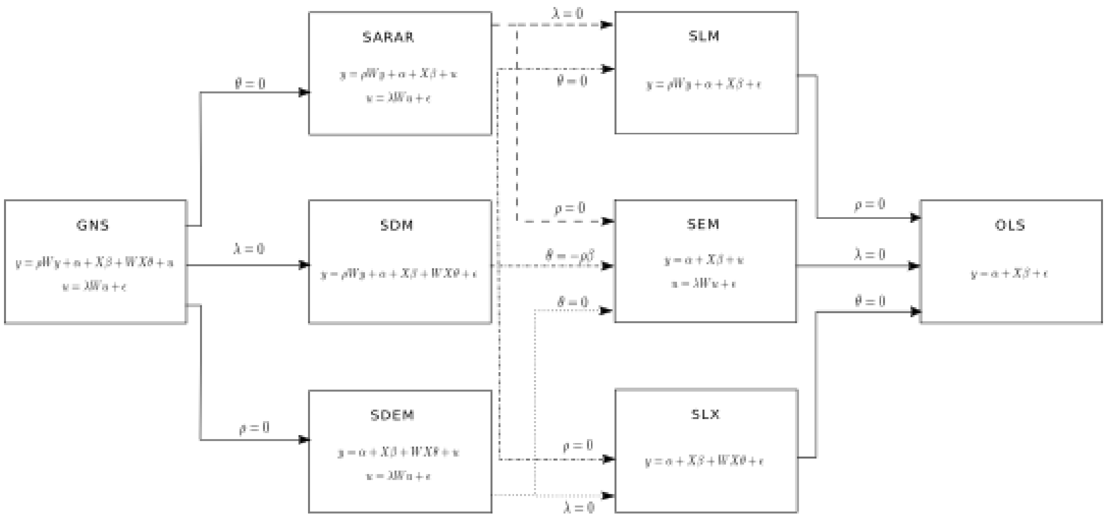

Figure 11 Relationships between spatial models
Source: Vega & Elhorst, 2015.

Within the models of spatial dependence, an important conceptual separation is between *spatial error* models and *spatial lag* models. Spatial error models look for spatial dependence in the model errors. In comparison, spatial lag models include a spatial lag of a variable as an explanatory variable. Although the distinction of models can fade away in higher-order models where both effects can be incorporated in the same model, the distinct interpretations of spatial errors and spatial lags still exist.

The implications of spatial dependence could change based on whether the dependence is a spatial error or a spatial lag. On one hand, spatial lags are associated with peer-effects, spatial process consequences or spill-overs. If, for instance, there is a behavioral process in which the prices are partly determined by the prices of the nearby observations, then there is a spatial lag of dependent variable expected to the extent the effect is not controlled by spatial dependence naturally embedded in the independent variables. Similarly, a spatial lag may also be a result of a spatial process. If, for instance, a spatial process such as a traffic caused by proximity to a point of interest, affects the prices, a spatial lag of the dependent variable (price) is also expected to be observed in the data. Comparatively, a spatial lag of independent variables can also be used as explanatory variables, as in an SLX, indicating a spatial spillover of the independent variables on the dependent variable. Figure 11 also assists in finding relevant models for observing desired effects. As seen in Figure 11, for spillover effects, where spatial lags of independent variables are expected to have an effect on the dependent variable, the models SLX, SDM, SDEM or GNS are appropriate. For peer-effects, where the spatial lags of the dependent variable are expected to have an effect on itself, SLM, SDM, SARAR and GNS models are appropriate. On the other hand, spatial errors have no such direct interpretations other than a spatial definition mismatch and they are relevant because the omitted spatial effects cause increased variation in residuals, thus making it harder to detect the changes in parameters of interest.

The consequences of spatial errors and spatial lags are thus different. Ignoring spatial errors in OLS result in an efficiency problem but does not result in an estimation problem. In contrast, ignoring spatial lags results in an estimation problem and the estimates become biased and inconsistent when spatial lags are not included. Therefore, ignoring spatial dependence risks at least a potentially wrong fit and standard errors in OLS and it can result in biased and inconsistent estimates as well.

Specifically in real estate pricing models, the analysis in Chapter 3 shows that the spatial regressions make some previously significant spatial fixed effects statistically insignificant and some previously insignificant building characteristics statistically significant, therefore change the inference of the pricing analysis. While these changes in the spatial regression coefficients and the regular hedonic regressions, and the non-linear effects of the variables discourage the use of the standard hedonic regression in house pricing in theory, Chapter 3 also presents practical effects of spatial dependence with standard hedonic regression.

CHAPTER 3

AN ANALYSIS OF SPATIAL DEPENDENCE

3.1 Abstract

Real estate properties are naturally space-fixed, therefore, a spatial dependence of prices is expected. When space-specific factors are not fully incorporated in pricing equations, spatial autocorrelation is likely to exist as happens in standard hedonic regression residuals. This results in problems in the interpretation of coefficients either due to inefficient estimators in simple models or to incomprehensibility in complex models. Thus identifying spatial factors in a domain is important and the focus of this study is to identify and test the determinants of spatial dependence in the housing market.

Using a novel data set, the study contributes to the literature as factors of spatial dependence are identified and tested in a metropolitan housing market and the literature is extended to an emerging market. Moreover, in this study, the spatial clustering of housing characteristics such as area, total number of floors and the building age is documented. The performance of the hedonic regression equations is evaluated when the determinants of spatial dependence are controlled for. Supporting results are found for the hypotheses that the construction process, shared social services and high-rise residential complexes cause spatial correlation. The findings show that spatial correlation is significantly reduced when the factors of spatial dependence and district level data is controlled for in the standard hedonic regression.

3.2 Introduction

The models for pricing real estate properties involve a variety of factors including spatial characteristics. This results in each single property having its own dynamics of pricing. Location is one factor, though, space does not always influence a property uniquely. The spatial effects can be property-specific, such as having a particular view; area-specific, for example, the proximity to public transportation; neighborhood-specific like the average income level of the residents; or it could be more general such as a major investment plan for the city.

The spatial factors result in spatial dependence, thus, clustering in property prices across space. If space-specific factors persist for some time, spatial autocorrelation is likely to exist in a standard hedonic regression for the cross-section of prices as space-specific factors will not be fully incorporated in pricing equations. Since spatial autocorrelation makes OLS estimator inefficient and omitting locational factors risks a bias, the use of standard hedonic regression might be problematic in inferring a causal model. To overcome the problem of inefficient estimators, several econometric models have been introduced over the last few decades. While using such models facilitates better use of the data, it increases the model complexity and complicates economic interpretation of spatial factors. Therefore, identifying the factors of spatial dependence is of interest; moreover, separating out the locational effects lays the groundwork for quantifying and predicting effects of particular locational factors on real estate prices.

In this study, the causes of spatial dependence are analyzed using a novel data set for the metropolitan city of Istanbul. Cross-sectional data used in this analysis consist of 25,219 observations of real estate listings for housing sales. First, a standard hedonic regression is run and the spatial autocorrelation in residuals are tested. Then, spatial analyses of residuals, covariates and subsamples are performed. The results find supporting evidence for the construction process, shared social services, and high-rise residential buildings to cause spatial autocorrelation in standard hedonic regressions. Particularly, building age and the total number of floors exhibit high spatial cross-correlation with the regression residuals. However, when all the spatial factors and district level data is controlled for, the performance of the hedonic regression analysis improves significantly. The performance improves even further when district level data is replaced with neighborhood level data.

This study contributes to the literature by focusing on the domain-related causes of spatial dependence in a metropolitan housing market. This is one of the first studies to examine the determinants of spatial dependence in housing prices. Moreover, the study is extended to an emerging market. Furthermore, using a novel data set including geolocation, this study has a comprehensive coverage of online listings and enables a more precise consideration of space for the analysis of house prices.

The remainder of this study is organized as follows. In Section 3.3, a review of the related literature is presented. In Section 3.4, the hypotheses and the methodology are outlined. Section 3.5 defines the data and Section 3.6 presents the analysis. Finally Section 3.7 presents a conclusion.

3.3 Literature review

3.3.1 Spatial autocorrelation measures: rests for spatial dependence in errors

3.3.1.1 Moran's I

Moran introduced the first spatial autocorrelation measure in 1950. The measure, later named as Moran's I, has become a widely-used measure for detecting spatial autocorrelation and it is defined as follows for a point sample:

$$Moran's\ I = \frac{N}{\sum_i \sum_j W_{ij}} \frac{\sum_i \sum_j W_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \qquad (30)$$

$$where\ N\ is\ the\ number\ of\ spatial\ units, x\ is\ the\ variable\ of\ interest$$

$$and\ W\ is\ a\ matrix\ of\ spatial\ weights$$

The measure is generalized and it's asymptotic properties are derived by Cliff and Ord in 1972, allowing its use in regression residuals. The way to use this test in real estate pricing context would be to check whether residuals are spatially correlated.

$$Moran's\ I = \frac{N}{\sum_i \sum_j W_{ij}} \frac{\hat{\varepsilon}^T (W)\, \hat{\varepsilon}}{\hat{\varepsilon}^T \hat{\varepsilon}} \qquad (31)$$

$$where\ N\ is\ the\ number\ of\ spatial\ units, \hat{\varepsilon}\ is\ estimated\ residuals\ vector$$

$$and\ W\ is\ a\ matrix\ of\ spatial\ weights$$

3.3.1.1.1 Statistical inference with Moran's I on regression residuals

Cliff and Ord (1972) allow approximate statistical inference on Moran's I statistic by deriving the moments of the measure, its sampling distribution and its asymptotic distribution. They analyze three ways of statistical inference based on the sample:

under normality assumptions of residuals, under random permutations and using Monte Carlo methods.

3.3.1.1.1.1 Inference under normality assumption

They show, based on Pitman–Koopman–Darmois theorem under the normality assumption of the residuals, that Moran's I statistic is asymptotically distributed normally with the condition that the standardized cumulatives of the statistic approach zero as N gets large. Thus, the inference is made using the asymptotic normal distribution of the statistic. Anselin (1988) presents a generalized version of the expected value and variance of Moran's I is for a general (non-standardized) spatial weight matrix as:

$$E[Moran's\ I] = \frac{N}{\sum_i \sum_j w_{ij}} \frac{tr(MW)}{N-K} \tag{32}$$

$$V[Moran's\ I] = \left(\frac{N}{\sum_i \sum_j w_{ij}}\right)^2 \frac{tr(MWMW') + tr(MW)^2 + (tr(MW))^2}{(N-K)(N-K+2)} - (E[Moran's\ I])^2 \tag{33}$$

$where\ N\ is\ the\ number\ of\ spatial\ units,$

$M\ is\ the\ projection\ matrix\ (I - X(X'X)^{-1}X', W\ is\ a\ matrix\ of\ general\ spatial\ weights,$

$and\ K\ is\ the\ rank\ of\ the\ regression$

Griffith (2010) later argues that the first two moments of Moran's I are robust to violations of normality assumption, even in the case of heteroscedasticity, when the sample size is large.

3.3.1.1.1.2 Inference under random permutations

Cliff and Ord (1972) also show that inference can be made by calculating Moran's I for random permutations of the observations on the defined space. This relaxes the assumption of i.i.d residuals, while it implicitly assumes lack of correlation among x_i's and reduces statistical power due reduced information. Regardless, permutation is not an alternative for Moran's I calculation from regression residuals. The reason is that the permutations of the error term is equivalent to permutations of the vector of dependent and independent variables and such permutations ignore the autocorrelation among the independent variables (A. D. Cliff & Ord, 1981).

3.3.1.1.1.3 Exact distribution of Moran's I

In addition to the asymptotic distribution of Moran's I provided by Cliff and Ord (1972), the exact distribution of Moran's I for regression residuals is given by Tiefelsdorf and Boots (1995) under the normality assumption of the residuals. Additionally, in the case of correlated residuals, upper and lower bounds for the distribution can be derived. This methodology is especially well suited for datasets with low number of observations, where the statistic is further from converging to its expected value and where the correlation between the residuals may have more dramatic effect on the inference. A drawback of this method is that it requires calculation of eigenvalues for the spatial structure, which is a computationally expensive procedure that requires $O(n^3)$ time complexity in practical benchmarks (Anderson et al., 1999). For computational concerns, a saddlepoint approximation of Moran's I distribution is suggested by Tiefelsdorf (2002), which still involves eigenvalue calculations therefore might be better suited for smaller data sets. Figure

12, by Tiefelsdorf and Boots (1995), shows a demonstration of the upper and lower bounds for a distribution. The convergence of upper and lower bounds in the figure suggests the robustness of the measure as sample size gets large.

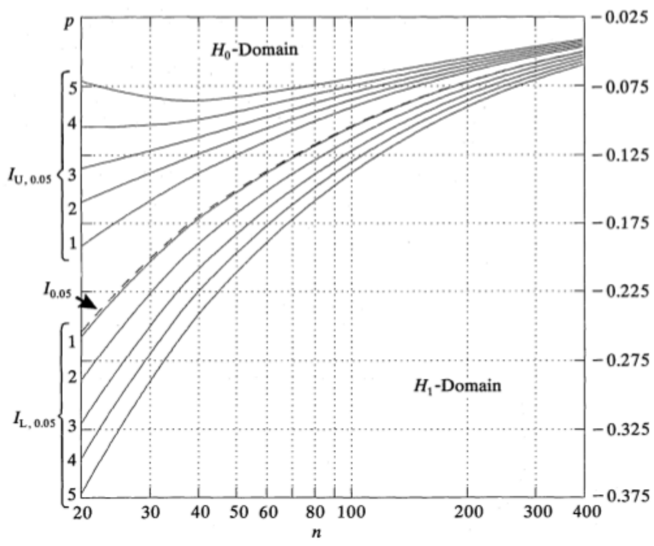

Figure 12 Demonstration of Moran’s I’s bounded distribution with respect to its exact distribution under one-sided test of regression residuals where space is defined as rings of hexagons

Source: Tiefelsdorf and Boots, 1995.

3.3.1.1.1.4 Null hypothesis of Moran’s I

In any case of inference method explained above, the null hypothesis for Moran's I statistic is that the attribute being analyzed is randomly distributed in the defined space. Thus, the rejection of the null hypothesis with a significant p-value using a two-sided test means that the spatial distribution of attributes is more spatially clustered or dispersed than a random spatial process would generate. It is also possible to use a one-sided test, such as in cases where negative spatial dependence is not expected, in which case the rejection of null hypothesis would mean the observed spatial distribution is more clustered than a theoretical non-spatial-dependent data generating process.

3.3.1.1.2 Defining spatial weights matrix

Defining the spatial structure is a crucial part measuring spatial autocorrelation (and spatial modeling in general) as it sets up the rules on how observations are affected over and through space. In fact, defining the spatial structure in measuring spatial autocorrelation is more influential to the results than selecting a spatial model (Bell & Bockstael, 2000).

In most spatial autocorrelation measures, including Moran's I, space is defined by an (n x n) spatial weights matrix (W) that shows the spatial relation between pairs of observations. Alternatively, specification by functional form without a spatial weights matrix is also possible. (Kelejian & Robinson, 1992). Common approaches in constructing a spatial weights matrix are to use binary indications or weighted measures of neighborhood relations that are based on somewhat arbitrary rules about shared borders and/or distances (e.g. k-th neighbor, distance-boundary weights, radial distances etc.). Such approaches are especially useful for defining space when observations under study are areal units such as districts of a city, where space relations are further complicated by artificial boundaries and the data on observations might be aggregated over such units. For point-type analysis involving micro agents, using distance-based weighting is more suited. The functional form such as power distances, exponential distances or radial distances and its parameters are still to be decided by the researcher.

A further issue in defining spatial weights matrix arises from the discontinuity in the parameter space. As Bell and Bockstael (2000) presents, $(I - \rho W)$ has to be positive and non-singular in a regression model to be defined with spatially autocorrelated residuals such as:

$$y = \beta X + \epsilon, \quad \epsilon = \rho W \epsilon + u \tag{33}$$

$$y = \beta X + (I - \rho W)^{-1} u \tag{34}$$

The eigenvalues of W, which changes with the sample, is directly related with where the parameter space is not defined. Therefore, row-standardization is a common practice in defining the spatial weights matrix W, where each row is normalized so that their sum equals to one. The row standardization ensures the problem to be defined over -1/$|\lambda_{min}|$ to 1 (Bell & Bockstael, 2000). Moreover, row-standardization also simplifies Moran's I such that the first term $\frac{N}{\sum_i \sum_j w_{ij}}$ necessary equals to 1 with a row-standardized W and Moran's I measure becomes (Arbia 2014):

$$Moran's\ I = \frac{\hat{\varepsilon}^T (W)\, \hat{\varepsilon}}{\hat{\varepsilon}^T \hat{\varepsilon}} \tag{35}$$

$N\ is\ the\ number\ of\ spatial\ units, \hat{\varepsilon}\ is\ estimated\ residuals\ vector$

$and\ W\ is\ a\ matrix\ of\ row\ standardized\ spatial\ weights$

The disadvantage of row standardization is that it might alter the interpretation of spatial effects, since it changes the defined spatial structure. For this reason, Anselin (1988) suggests that row standardization should not be done automatically. According to him, although it is a common practice, it is not a statistical necessity (L. Anselin, 1988).

Another concern in defining a weight matrix is that no particular subset of units should dominate the spatial structure for the expected asymptotic distribution of the measure to converge to a normal distribution. Cliff and Ord (1972) gives the following sufficient but not necessary condition for such a weight matrix:

(i) $(W_{ij} + W_{ji}) \,/\, W$ are all $O(N^{-1})$ or less;

(ii) For each unit, only a finite number of $(W_{ij} + W_{ji}) \,/\, W$ are $O(N^{-1})$

If both of these conditions are satisfied, then the standardized cumulatives of Moran's I goes to zero as N gets large.

3.3.1.1.3 Criticism and modifications to Moran's I

The null hypothesis of Moran's I test is that there is not any spatial autocorrelation present in the variable. Rejecting the null hypothesis in a regression context therefore means there is a spatial dependence in the error terms; however, there was no alternative hypothesis defined that would suggest a specific type of spatial dependence earlier. Nevertheless, Burridge (1980) showed that Moran's I measure is equivalent to a Lagrange Multiplier test with a spatial error model (SEM) as an alternative. Therefore, conclusions from Moran's I will be the same as the ones from an LM test with SEM as an alternative (Arbia, 2014).

Another addition to Moran's I measure was from Kelejian and Prucha (2001). They generalize the normalization factor ($\frac{N}{\sum_i \sum_j W_{ij} \;\; \hat{\varepsilon}^T \hat{\varepsilon}}$) in Moran's I such that it takes a different form based on specific alternative hypothesis.

Even though, the conclusions from Moran's I will be similar to LM-SEM, Moran's I is not an estimator of the autocorrelation coefficient λ of SEM (L. Anselin, 1988; Li et al., 2007). Li, Calder and Cressie (2007) argue that Moran's I only estimates λ well around 0.

3.3.1.2 Alternative measures of spatial dependence

3.3.1.2.1 Measures of spatial dependence

In addition to Moran's I, measures of spatial dependence suggested in the literature include Mantel's test (Mantel, 1967), Geary's contiguity test (Geary, 1954), Getis-Ord General G (Getis & Ord, 1992), K-R test (Kelejian & Robinson, 1992) and LM based tests (LM-ERR: Burridge, 1980; LM-EL: Luc Anselin et al., 1996).

Geary's C is shown to be asymptotically equivalent to Moran's I (Chun & Griffith, 2013). LM-based tests have specific alternative hypothesis indicating the existence of alternative hypothesis in case of rejecting the null, as opposed to Moran's I which has no specific alternative hypothesis. The alternative hypothesis for LM-ERR is the SEM model and the alternative hypothesis for LM-EL is the SLM model. Importantly, Burridge (1980) showed that LM-ERR test is the square of Moran's I and therefore the conclusions from both tests are the same. The alternatives indicated by specific LM test do not reject the other spatial model, therefore robust versions of those tests are suggested. The K-R test, unlike other tests, measures spatial dependence by functionally specifying the variance-covariance matrix. An additional measure is suggested by Li et al. (2007) for specifically estimating the coefficient of SEM, as they argue that the coefficient of SEM is usually what researchers mean by spatial dependence.

3.3.1.2.2 Local measures

Another set of spatial dependence measures are local measures, where the statistic indicates the spatial dependence based on a specific observation and how similar the values of other observations are with respect to space in between. Anselin (1995) described LISA (local indicators of spatial association) statistics, which calculate Moran's I for each unit:

$$Local\ Moran's\ I_i = z_i \sum_j (w_{ij} z_j) \tag{36}$$

$where\ z_i\ are\ standardized\ and\ centralized\ variables$

$and\ W\ is\ a\ matrix\ of\ spatial\ weights$

Geary’s C measure can be calculated locally similarly as an alternative local measure. An important feature of Local Moran’s I statistic is that the sum of all Local I’s gives the global Moran’s I.

$$\sum_i Local\ Moran's\ I_i = Moran's\ I \tag{37}$$

Therefore, local measures are used to detect whether local spatial dependencies exist while it might be the case that there is no global spatial dependency observed. It is usually used in cluster maps, significance maps and scatter plots. The significances of local Moran’s I statistics are judged by permutations conditional upon the unit in question.

3.3.1.2.3 Spatial cross-correlation

While spatial autocorrelation is observed when a variable is related to itself through a spatial definition, spatial dependence might be observed across variables as well. In other words, multiple variables could be related spatially, while they may be uncorrelated otherwise. Wartenberg (1985) extends spatial dependency to a multivariate relationship:

$$M = Z^T W Z \tag{38}$$

$where\ Z\ is\ a\ n\ by\ m\ matrix\ of\ standardized\ and\ centralized\ n\ variables,$

$Z^T is\ the\ transpose\ of\ Z, and\ W\ is\ a\ matrix\ of\ spatial\ weights.$

$M\ denotes\ the\ spatial\ (cross)\ correlation\ matrix.$

Following Wartenberg, a spatial cross-correlation coefficient is given by Reich, Czaplewski, & Bechtold (1994) for two variables as:

$$Moran's\ Bivariate\ I_{yz} = \frac{1}{\sum_i \sum_j W_{ij}} \frac{\sum_{\substack{i \\ i \neq j}} \sum_j W_{ij} y_i z_i}{\sqrt{var(y)\ var(z)}} \tag{39}$$

$where\ y\ and\ z\ are\ standardized\ and\ centralized\ variables,$

$and\ W\ is\ a\ matrix\ of\ spatial\ weights$

Czaplewski and Reich (1993) gives the expected value and variance for this coefficient is given as follows, while stating that Moran's Bivariate I is distributed normally for samples larger than 40 observations:

$$E[I_{yz}] = \frac{-\rho_{YZ}}{n-1} \tag{40}$$

$$var(I_{YZ}) = \frac{\begin{array}{r}(\frac{m_{YZ}^2 n}{m_{Y^2} m_{Z^2}}[2(W^2 - S_2 + S_1) + \frac{S_1 - S_2}{2}(n-3) + \frac{S_1}{2}(n-2)(n-3)] \\ + \frac{-m_{Y^2Z^2}}{m_{Y^2} m_{Z^2}}[6(W^2 - S_2 + S_1) + (4S_1 - 2S_2)(n-3) + S_1(n-2)(n-3)] \\ +n[(W^2 - S_2 + S_1) + \frac{S_1 - S_2}{2}(n-3) + \frac{S_1}{2}(n-2)(n-3)])\end{array}}{(n-1)(n-2)(n-3)W^2} - E[I_{yz}] \tag{41}$$

where $m_Y^2 = var(y)$, $m_Z^2 = var(z), m_{YZ} = cov(y,z), m_{Y^2Z^2} = \sum_{i=1}^n y_i^2 z_i^2/n, S_1 = \sum_{i=1}^n \sum_{j=1}^n W_{ij}^2, S_2 = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n w_{ij} w_{jk}$ and ρ_{YZ} is the linear correlation coefficent.

As with Local Moran's I, Moran's Bivariate I can be calculated locally for each unit:

$$Local\ Moran's\ Bivariate\ I_{yz_i} = \frac{1}{\sum_i \sum_j W_{ij}} \frac{\sum_j W_{ij} y_i z_i}{\sqrt{var(y)\,var(z)}} \tag{42}$$

where y and z are standardized and centralized variables,

and W is a matrix of spatial weights

Chen (2015) similarly defines SCI (Spatial Correlation Index) and LSCI measures (Local Spatial Correlation Index) for assessing spatial relationships between two variables:

$$SCI_{yz} = y^T W z \tag{43}$$

$$LSCI_{yz_i} = y_i \sum_j W_{ij} z_j \tag{44}$$

where y and z are standardized and centralized variables,

and W is a unitized matrix of spatial weights

Note that while the number SCI_{yz} equals to SCI_{zy}, this is not necessarily true for local measures. In other words, the vector $LSCI_{yz}$ is in general not the same as the vector $LSCI_{zy}$.

Additionally, Chen (2015) incorporates space to regression analysis to judge causality between two variables:

$$nWz \ = \beta y \tag{45}$$

$where\ y\ and\ z\ are\ standardized\ and\ centralized\ variables,$

$W\ is\ a\ unitized\ matrix\ of\ spatial\ weights,$

$\beta\ denotes\ the\ regression\ coefficient,$

$and\ n\ denotes\ the\ number\ of\ observations$

Note that the regression has no intercept and the relationship is judged by F statistic in a regular manner.

3.3.2 Spatial autocorrelation in real estate pricing

Dubin (1988) provides a maximum likelihood estimation technique for regression coefficients using covariance specification when spatial autocorrelation is present. This paper is also an early paper where housing market is used as an example of spatially autocorrelated series, estimating a hedonic regression with cross-sectional data. Can (1992) utilizes spatial expansion techniques, allowing housing attributes to vary over space by adding interaction variables, to incorporate spatial dependence in housing price models. Later in 1990's this strand of research gets traction. Pace, Barry & Sirmans (1998) describes modeling neighborhood relations for weight matrices representing space (as opposed to covariance matrices). Dubin (1998) compares these two methods, weight matrix and covariance matrix approaches, and reports that two methods assume different spatial relation and any spatial model is better than no model when spatial autocorrelation is likely to be present. Basu & Thibodeau (1998) reports strong presence of spatial autocorrelation in residuals among single-family property transactions in Texas, using covariance matrix approach.

Due to the problems in optimality of OLS when there is autocorrelation, Kelejian & Prucha (1998) propose, in their influential paper, a feasible GS2LSL (generalized spatial two-stage least squares) procedure as an alternative to ML estimations. This procedure estimates linear regression coefficients for models that has a spatially lagged dependent variable and a spatially autoregressive residuals. This procedure becomes a widely-used procedure in regressing spatially dependent models, given its conceptual simplicity. Dubin, Pace, & Thibodeau (1999) also describe procedures for running regression estimations in order to deal with inefficient OLS estimators when there is autocorrelation. Specifically, they explain two contemporary methods, EGLS (estimated generalized least squares) and ML (maximum likelihood) that have been used in spatial regressions. They provide a 3-step calculation of EGLS. First, data is fitted using OLS. Second, separating distances are divided in ranges and correlations between residuals for each range are calculated. Third step is to use the correlations to improve OLS prediction (kriging).

Bell & Bockstael (Bell & Bockstael, 2000) applies GS2LSL procedure to Maryland housing market and they criticize that the weight matrix approach changes the observed economic relationship by changing the observed spatial structure. Gillen, Thibodeau, & Wachter (2001) point out the anisotropy in the autocorrelations. Wilhelmsson (2002) applies a number of spatial models to real estate data and confirms that spatial autocorrelation is present, OLS estimates could be biased and spatial models explain more of a variation, while noting that the spatial model choice does affect economic interpretation.

More recently, Bitter, Mulligan, & Dall'erba (2007) apply GWR to the real estate context and report that GWR outperforms spatial expansion techniques. Tu, Sun, & Yu (2007) uses spatial autocorrelations to identify submarkets in urban housing. Huang, Wu, & Berry (2010) report a positive performance of GWR based models in the context of house pricing in Calgary of Canada. Bourassa, Cantoni, & Hoesli (2010) compares geostatistical methods to other methods in the case of a housing market of a county in Kentucky. They favor geostatistical methods with a consideration of submarkets to other methods. Osland (2010) uses several of the spatial methods to improve house pricing model specification.

3.3.3 Possible causes of spatial autocorrelation in real estate

As with any regression, misspecification is a possible source of autocorrelation. If, for instance, an important variable is omitted and the omitted variable is not distributed randomly across space, then the residuals will be biased across space as well. Consider theoretical grouping of factors affecting housing prices by Basu and Thibodeau (1998):

$$Price = f(L; S; N; A; E; R; t) \qquad (46)$$

$where\ L\ denotes\ lot\ characteristics,$

$S\ denotes\ structural\ characteristics\ of\ the\ building,$

$N\ denotes\ neighborhood\ characteristics,$

A denotes accessibility to the housing,

E denotes proximity externalities,

R denotes regulations regarding the use of land,

t denotes the time period that the housing data is collected,

Of all groups of factors affecting price only data collection time and building characteristics is not directly affected by space. Land characteristics, neighborhood characteristics, accessibility, proximity externalities and regulations are directly related to the specific location of the housing. Thus, any missing variable in these theoretical groups is a likely candidate of spatial autocorrelation statistically.

Gillen, Thibodeau, & Wachter (2001) points out the spatial dependence in the building process of housing. Spatially close buildings are usually constructed at similar time periods and thus are similar in terms of housing area, age and design features. The authors also claim that spatially close houses share social services such as services provided by fire departments, police stations, health institutions, and transportation services. For this reason, the effect of such services on the price is similar in spatially close buildings.

Sun, Tu, & Yu (2005) report that there is a building dependence in addition to neighborhood dependence for multi-floor residential houses. All housing units in a multi-floor residence building have the same location, therefore they require an update in models based on neighborhood effects.

Wiltshire (1996) points out the possible effect of interlinked personnel undertaking the valuation process on spatially close properties in case of valuation-based data sets.

Dunse, Jones, Orr, & Tarbet (1998) state market inertia, which might cause a smoothing in prices across properties with proximity, is a possible reason for spatial autocorrelation.

3.3.4 Real estate market pricing in Istanbul

Ozus, Dokmeci, Kiroglu, & Egdemir (2007) analyze Istanbul real estate market using 6 separate regressions at a district level and having sub-districts as dummy variables in those regressions. They state that sub-markets vary in their coefficients as well as significant variables. As for the metropolitan level, they report that sub-market, area and sea view affect housing prices most. Keskin (2008) makes use of a survey data containing socioeconomic and neighborhood characteristics in addition to housing data. She uses a single hedonic regression, including locational factors such as travel time to job & schools, school satisfaction, health service satisfaction, earthquake risk and continent. She reports that living area size, being in a low floor building, being in a secured site (with swimming pool and garage), and age of the building affect prices the most. Kaya (2012) undertakes a comprehensive study for detecting house price changes by using a different set of data where prices are determined by expert valuations. She reports cross-sectional hedonic results for Istanbul with a 0.59 to 0.72 R2, depending on the period of analysis. Her locational factors in those regressions are dummy variables for districts.

Koramaz and Dokmeci (2012), on the other hand, tries to explain house prices using only location variables in a hedonic regression. They report a similar R2 value to Kaya’s (2012) using only the following variables in interaction with area: distances to CBD, sub-centre, transportation arteries and coast, density and whether in the Bosphorus coastal region. However, their residuals are highly spatially dependent.

Even though there are a number of studies that looked into house pricing in Istanbul with locational variables, a very few of them included a spatial dependence model. Keskin (2010) utilizes a two level hierarchical regression methodology using Istanbul housing data in order to overcome problems caused by spatial autocorrelation. Bekar and Akay (2014) applies two-staged spatial quantile regression method, a method proposed by Kim and Muller (2004) that runs regressions with a spatially-lagged variable in quantiles. The method is essentially similar to the weighted matrix method in regards to modeling spatial dependence. However, the authors do not put emphasis on how they formed the spatial model.

To the author's best knowledge, Uyar and Yayla (2016) gives the first and only study that incorporates a discussion of spatial modeling in the context of pricing housing properties in Istanbul. They apply a spatial autoregressive model, a spatial error model with weight matrices and a spatial Durbin model and present a comparison to a regular OLS estimation. According to their results, all the spatial models give a better fit than regular OLS and the spatial Durbin model performs slightly better than the other two spatial models.

Table 6. Summary of Selected Studies on Istanbul Housing Market

Authors	Focus	Data			
		Source	Time-Period	Number of Observations	Scope
Ozus et al. 2007	Exploring price differences across 6 districts in Istanbul	Realtors	1997, Summer Period	1468 observations	Istanbul, picked from 26 sub-districts
Keskin, 2008	Exploring price differences across 348 submarkets in Istanbul	(a) 2 major Realtor Agencies (b) Survey by Istanbul Greater Municipality	(a) 2006, November and April	(a) 2175 transactions (b) 1517 responses	Istanbul, 32 districts, 946 neighborhoods
Kaya, 2012	Identifying house price changes across Turkey	Expert Valuation Reports for Mortgage Applications	2010.12 - 2012.06	487027 reports	Turkey, Including Istanbul and 26 districts
Uyar and Yayla, 2016	Applying Spatial Models to Istanbul Real Estate Market	Not reported	2013, October to December	2797 observations	Istanbul, sampled in a stratified manner across 39 districts
Morali, 2020 (this study)	Exploring spatial dependencies in housing market	(a) A major online marketplace for realtors (HurriyetEmlak. Com) (b) Primary data collection on significant locations	2017-10-15	25219 observations	Istanbul, housing market, 39 districts, 514 neighborhoods

3.4 Hypotheses and methodology

3.4.1 Hypotheses: causes of spatial dependence

Spatial autocorrelation is a technical issue regarding regression residuals, indicating that there is unused information left in the data and that the prediction and inferences might be misleading. How to improve the statistical procedures is extensively studied in the literature, providing us a variety of spatial models.

Since spatial autocorrelation is essentially a mismatch between the statistical model and the data generating process, another perspective is on spatial autocorrelation is that it gives insight on how the data generating processes do not match the models commonly used elsewhere (in addition to how the models do not match the data generating processes). In other words, it presents an opportunity on elaborating how real estate processes might have occurred.

Findings in the literature suggest that regression residuals exhibit spatial dependence (Can, 1992; R. A. Dubin, 1988; Pace et al., 1998; Wilhelmsson, 2002). The spatial dependence in regression residuals is tested using spatial autocorrelation measures and used as a point of reference for further analysis. Hence, the first hypothesis (H1) is as follows:

Hypothesis 1: There is spatial dependence in the standard hedonic regression residuals.

One cause of spatial dependence is the construction process of housing. It has been pointed out that spatially close buildings are usually constructed at similar time periods, thus, are similar in building age. Moreover, these properties share similar design features such as housing area and number of bedrooms (Basu and Thibodeau, 1998; Gillen, Thibodeau, & Wachter, 2001). Still, there might be uncontrolled variables in the regression equation such as the architectural design or there can be some nonlinear interactions, thus, the construction process causes spatial autocorrelation in residuals. Hence, the second hypothesis (H2) is as follows:

Hypothesis 2: Construction process is a determinant of spatial autocorrelation.

It has been suggested that spatially close properties share social services such as fire departments, police stations, health institutions, and transportation services (Dubin, 1988; Basu and Thibodeau, 1998; Gillen, Thibodeau, & Wachter, 2001). Consequently, the effect of such services on the price is similar in spatially close buildings and causes spatial autocorrelation if not adequately considered in a regular regression. Hence, the third hypothesis (H3) is as follows:

Hypothesis 3: Shared social services is a determinant of spatial autocorrelation.

Another important factor is building characteristics such as being a high-rise residential complex. For example Sun, Tu, & Yu (2005) report that there is a building dependence in addition to neighborhood dependence for multi-unit residential houses. All housing units in a multi-unit residence building have the same location, therefore they require an update in models based on neighborhood effects. High-rise residential units may result in spatial autocorrelation as they include a higher density of very similar dwellings at a particular location and, consequently, such residences tend to try to differentiate themselves from other buildings by providing security gates, social services and a community feeling. Hence, the fourth hypothesis (H4) is as follows:

Hypothesis 4: High-rise residential complexes are a determinant of spatial autocorrelation.

3.4.2 Methodology

As a general approach in this study, prices are regressed on attributes of the housing hedonically and spatial factors affecting house prices are analyzed. The sample size is 25,219. There are 39 provinces and 517 neighborhoods in Istanbul and geographic coordinates of the housing units are used in locational variables.

Empirical hedonic house price specification used is of the following general form, where the attributes are conceptually grouped into four for convenience:

$$ln(p_i) = B_i + F_i + A_i + L_i + \varepsilon_i \quad (47)$$

$where\ p\ denotes\ house\ prices$

$B\ is\ a\ set\ of\ variables\ regarding\ structural\ characteristics\ of\ the\ building$

$F\ is\ a\ set\ of\ variables\ regarding\ flat\ characteristics$

$A\ denotes\ the\ dummy\ variables\ for\ fixed\ effects\ in\ the\ administrative\ areas$

$L\ is\ a\ set\ of\ variables\ based\ on\ specific\ geolocation\ of\ the\ building$

Moran’s I statistic is used to measure the spatial dependence in the error terms of the regression. This selection of measure is due a number of reasons. First, it is by far the most commonly used statistic for spatial autocorrelation (Kelejian & Prucha, 2001). It is one of the first measures that was introduced in the mid 20th century and it became popular over the years. It has been applied to a variety of circumstances in different fields. Its behavior and properties are well studied in the literature, including its use in regression residuals. Statistical inferences are usually made with a normality assumption of residuals in its use in regression. Since the sample size is large in this research, the normality assumption is easier to make. Moreover, it has desirable asymptotic properties, especially for large samples. Its asymptotic distribution is normally distributed and based on Tiefelsdorf and Boots' (1995) work, the serial correlation in residuals seems to be less of a problem as sample size gets large. Additionally, Griffith (2010) argues that its mean and variance are robust for samples over a hundred in size even in the case of heteroscedasticity. Therefore, the normality assumption is used for the statistical inference method. Exact calculations are not preferred in this research due calculations' computational complexity and asymptotic equivalence to normal distribution.

There are a number of ways to implement Moran's I, other than the original specification. For instance, Kelejian & Prucha (2001) generalized it so that it's expected variance is based on a specific alternative. In this research, Anselin's (1988) specification for general weights is followed. The software implementations in this study of spatial methods are self-developed in R environment with an objective of calculation efficiency, and their results are tested with the spdep library where applicable.

For the hedonic regression, OLS estimation is used due intuitive interpretation of coefficients. Robust methods are not used because robust methods in general downweights the observations further to the mean in minimizing the errors, effectively changing the error definitions. The focus of this study, on the other hand, is to expose any spatial dependence that the data has such that possible model errors could be detected.

Given normality assumption, the Z-scores and p-values for statistical inference regarding Moran's I measure are calculated in a standard manner:

$$Z(MI) = \frac{MI - E(MI)}{V(MI)} \tag{48}$$

As one-sided test is more appropriate in hedonic house price regressions, the p-values at 95% confidence level for Moran's I measure become:

$$p_{MI} = prob(Z(MI) > 1.645) < 0.05 \quad (49)$$

When comparing spatial autocorrelation strength of a list of regression residuals with another, a test is performed against the probability that the difference is observed by chance. In this test, the normality assumption previously relied on with an additional independence assumption of both distributions is needed. The Z-scores in this case are calculated as follows:

$$\Delta MI = MI_{u_1} - MI_{u_2} \quad (50)$$

$$Z(\Delta MI) = \frac{\Delta MI}{max(V(MI_{u_1}),V(MI_{u_2}))} \quad (51)$$

In this comparison, deviations in both tails are plausible, therefore the p-values at 95% confidence level for comparisons become:

$$p_{\Delta MI} = prob(Z(\Delta MI) > |1.96|) < 0.05 \quad (52)$$

Spatial weights are calculated based on the (inverse) distances between observations within a distance band with a small correction term of 1 meter. Units further apart then 30 kilometers are considered to be not related. The distances are calculated as great-circle distances in meters based on WGS84 ellipsoid parameters.

$$W_{ij} = 1 / (D_{ij} + 1), if\ D_{ij} <= 30000 \quad (53)$$

$$W_{ij} = 0, if\ D_{ij} > 30000 \quad (54)$$

where $D_{ij}$is the distance between units i and j.

The spatial weights matrix is globally standardized, keeping the spatial inference in place.

$$w_{ij} = \frac{w_{ij}}{\sum w_{ij}} \tag{55}$$

In calculating spatial relations of different variables, self implementation of spatial cross-correlation measures explained in section 3.5.1.2.3 is followed.

H1 is tested by using Moran's I measure and the results are further examined by using additional measures and tools such as variograms and likelihood tests. H2 is tested by analyzing spatial cross-correlations between the baseline regression and proxies of design features. H3 is tested by including a new set of locational variables to the baseline regression model. The new variables contain the distances between the social service provider and the specific geolocation of the unit. H4 is tested by analyzing spatial cross-correlations and comparing measures in different subsamples to bootstrapped defaults. All hypotheses are also examined using alternative methods as robustness checks.

3.5 Data and the domain

Istanbul has a population of 15.0 million[13] living on two continents, with the Bosphorus in between. It has 39 districts and 964 neighborhoods located on a 5461 km2[14] of land (See Figure 13 for a division of the administrative units on map.). In

[13] Population in the year of data, 2017. Statistics from TURKSTAT.

[14] Statistics from Türkiye Mülki İdare Bölümleri Haritası, Harita Genel Müdürlüğü.

2017, it had received a 1.5% population increase and 238,383 housing units were sold.[15]

The data set used is a cross-sectional data set of 25,219 observations of housing listings for sale across Istanbul as of October 15, 2017. This data set is provided by HurriyetEmlak.com. HurriyetEmlak.com is a well-known marketplace

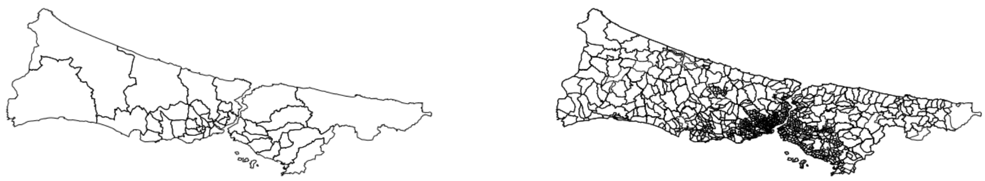

Figure 13 Districts and neighborhoods of Istanbul on a map

for real estate listings in Turkey as it is one of the two leading online portals and has a large coverage of the market across Turkey. The data set includes all listings in Istanbul that the provider had at a cross-section; therefore, data is considered to be robust to sampling bias and inflationary time effects during data collection period. The number of observations, 25219, is comparable to the average number of monthly sales in Istanbul, 21,024.[16] Regardless, as explained later in this section, the study focuses on housing that has no legality issues; therefore, the results could only be generalized on regular but investment grade housing properties.

[15] Statistics from TURKSTAT.

[16] Author's calculation based on statistics provided by TURKSTAT, indicating average of last 3 months of housing sales in Istanbul from the date of the data set.

Each observation consists of an attribute set over 20 variables including features like asking price, the area in square meters, property types, date added, building age, number of floors in the building, and number of rooms, bedrooms and bathrooms. (See Appendix A for the full details of the variables.) Main housing property types are flats, luxury residences and single houses. The location of an observation is given by the hierarchical administrative areas that the property belongs to in neighborhood detail and by geolocation of the property specifically. Figure 14 shows a mapping of housing sale listings in Istanbul.

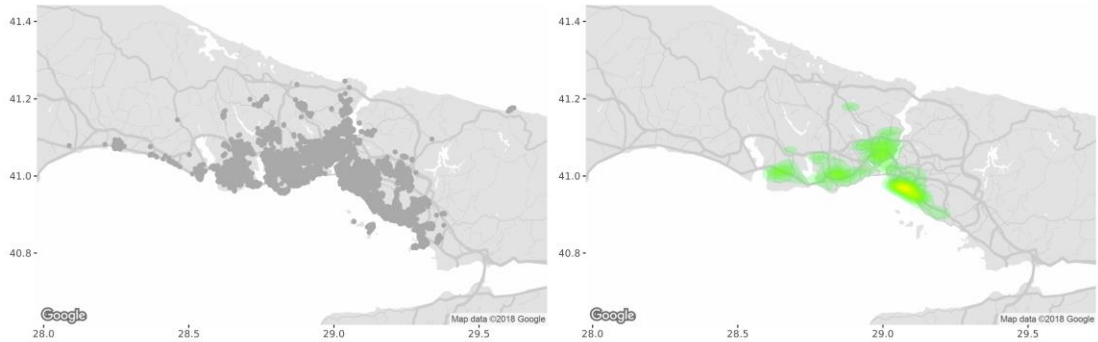

Figure 14 Locations of the units on map and the density

The data has gone through a multi-level screening process. The data was initially screened and approved to be active by the provider. Additionally, observations with implausible variables (e.g. higher buildings than the tallest known building, inconsistent in some variables, too high or too low prices) are removed, which are most likely to be data entry errors, non-standard offerings or inaccurate listings. Duplicate entries are also removed programmatically based on their location and explanatory variables. A small portion of deleted duplicates might be legitimate listings (as for instance in a possible case of a newly-built multi-unit condominium project); regardless, this is considered as a plausible trade-off as they constitute a very small portion relative to real duplicates if they exist and deleting these listings is likely to decrease spatial dependence if it has any effect on the results. Therefore, allowing them in the dataset would have amplified the inferences.

Istanbul is a city that has seen a rapid increase in population over the last decades and there has been a lack of long-term planning.[17] Consequently, the housing market suffers from legalities such as registration and permits.[18] For this reason, this research is restricted to “creditworthy” properties, which indicate whether the banks give a mortgage loan for those properties. In practice, Turkish banks require the property to have proper building registration and construction permits to lend a mortgage credit. Properties that are creditworthy but do not have residence permit are also excluded, which is stricter than the bank’s credit criteria as banks can provide mortgage loans for newly built houses without an issued residence permit.

A significant characteristic of the data set is that the prices are asking prices, which are expected to differ from transaction prices due to some margin of negotiation. The research questions in this study are of comparative nature, therefore exact valuation are not of primary interest. If, however, the average margin of negotiation is different at some regions, then spatially clustered error terms might exist. One reason for why the average margin of negotiation might be different in some regions might be the liquidity of properties. The literature suggests that the

---

[17] There has been 1667 zone plan changes in Istanbul between 2004-2006, while such plans should reflect long-term policies. (Turkish Court of Accounts Performance Audit Reports, 2008)

[18] Estimates based on expert opinion for buildings with legality issues goes up to include 70% of the buildings in Istanbul. (Turkish Court of Accounts Performance Audit Reports, 2008). The situation might have been improved over the recent years. Also note that this not the same as 70% of the population living in buildings with legality issues since multi-floor buildings are more likely to have proper permissions.

more illiquid properties tend to have higher times on the market (TOM) (Forgey et al., 1996; Kluger & Miller, 1990). For this reason, the listings that are older than 2 months[19] are excluded, downweighting the effect of liquidity and thus reducing the difference in margins of negotiations in different areas.

Lastly, this research is limited to multi-unit housing, excluding single family detached houses, townhouses or whole buildings for sale from the dataset. This is due to the fact that the explanatory variables might have different meaning when there is full ownership of the land. Moreover, this lack of economic meaning may also result in econometric problems as well. For instance, the "floor at" variable makes little sense in the case of full-ownership of land, therefore it has to be set to a specific value, which in turn causes a singularity issue between respective dummy variables. The area of the listing, for another instance, might indicate the area of the land in case of an old building sold for reconstruction, the area of a floor of the building, or the area of sum of all floors in the building. Such distinction is not standardized in the listings and there was no plausible way of distinguishing it. To avoid such problems, this research is restricted to buildings with multiple housing units.

This study also makes use of significant locations, such as public transportation stops and hospitals, schools or universities. This data is part of a primary data set that is collected by the author from respective public sources.

[19] The selection of 2 months is not arbitrary. It is the longest time allowed for a listing to be active at the marketplace without a renewal action.

3.6 Analysis

3.6.1 Spatial dependence in price

In order to establish a baseline for the analysis, the dependence of house prices on location is presented before going into detail on the dependence of regression residuals.

House prices in Istanbul are heavily dependent on space as one would expect and a visual plot demonstrates the common knowledge that the Bosphorus is deemed valuable in the city. Figure 15 is a map on which the data set is geoplotted with gradient coloring with respect to the unit's log price.

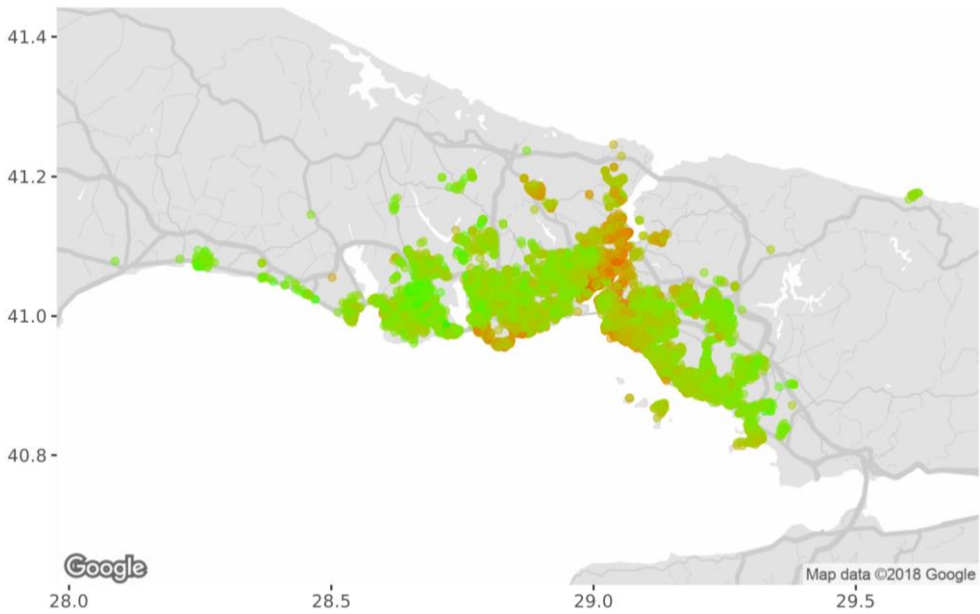

Figure 15 Spatial distribution of the data set observations on Istanbul map, gradiently colored by their log price levels

The visual inspection of Figure 15 shows that even though there are houses with high prices in most districts (seen as pale red dots on the map) the majority of the higher price units are concentrated along the strait. The map also shows how the housing units are distributed across the city. The density of housing and thus the population is concentrated on some regions and there are few local centers that are outside of the main clusters. From a non-spatial point of view, the prices show a fat tailed distribution, where there is more difference between high and average priced housing units than the normal distribution would have implied.

The spatial price dependence is analytically measured using Moran's I for point analysis and using empirical variogram. Note that the calculations of Moran's I for price points are slightly different from the calculations of Moran's I for regression residuals that are in the upcoming sections.

Table 7. Moran's I of Prices, Inference by Randomization

	Point MI	Expected MI	Variance MI	Z Score	P Value	n
Price	0.1315	-3.965e-05	8.87e-07	139.6	0	25,219

Moran's I value for the log of prices shows a strong spatial dependence in prices. The p-value of the test indicates that the observed Moran's I measure for the price points are significantly different from the expected Moran's I value considering the variance by rejecting the hypothesis of no spatial dependence for this data set using the normality assumption.

A variogram adds additional information to our understanding by expressing the variation with respect to distance. Empirical variogram plot below shows the empirical estimation of variogram by $2\gamma(h) = \frac{1}{N(h)} \sum_h [P_{s+h} - P_s]^2$. It indicates the variance of price difference between houses as a function of their distances. Figure 16 shows the empirical variogram of log of prices in the dataset.

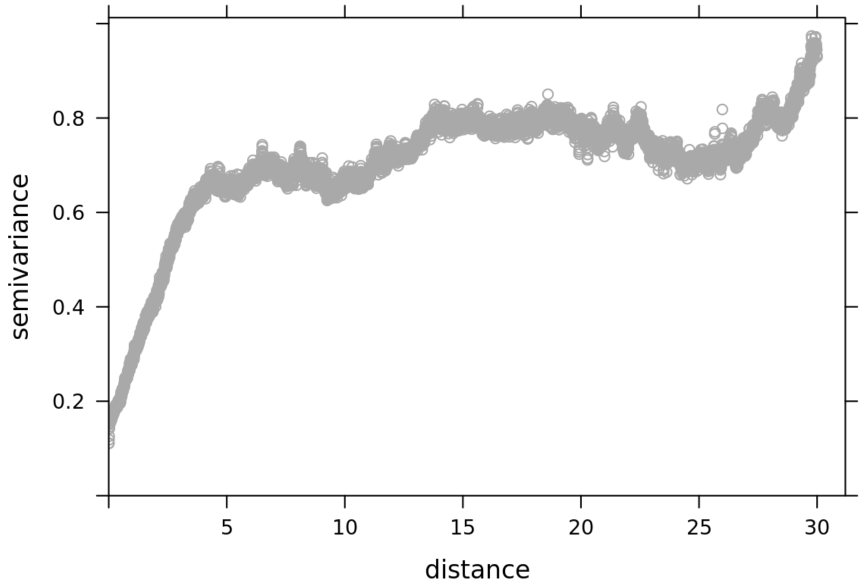

Figure 16 Empirical variogram of log prices, 5 meters bin-width, sparse bins (n < 30) are deleted, distance are in kilometers

The variogram in Figure 16 shows clear spatial dependence in prices. The range of the variogram seems to be around 5 kilometers. After 30 km's the variogram structure becomes unstable, which is likely to be due the geographics of the city. Since the N-S dimension of the city is bounded by the sea and parts of the city are linked over the sea in the E-W dimension at large distances, discontinuity of annuli increases for most points of the map in the variogram estimation as the distance increases over 30 kilometers. Note that due to the specific geolocation of the units and the high number of samples, it becomes possible to select bin-width as small as 5 meters. Variogram fits are omitted since the fitted values do not seem to be robust to changes in parameters of variogram and fitting method. Regardless, this is a problem about the fits, not about the variograms as the empirical variogram can go as precise as non-sparse 5 meters wide bins.

3.6.2 Spatial dependence in independent variables

In addition to the dependent variable (log of prices), the independent variables of the hedonic regression are also dependent on space. This is true not only for binary district variables that indicate the administrative areas that the apartment belongs to or for locational variables that indicate the distances to local centers but also for apartment characteristics such as number of bedrooms. Table 8 shows the spatial dependence of a set of selected variables.

Table 8. Moran's I of Independent Variables, Inference by Randomization

	Point MI	Expected MI	Variance MI	Z Score	P Value	n
Price	0.1315	-0.0000	0.00000	139.60	0.00	25,219
Age	0.1507	-0.0000	0.00000	159.86	0.00	25,219
Area	0.0831	-0.0000	0.00000	88.13	0.00	25,219
# of Bathrooms	0.1057	-0.0000	0.00000	112.16	0.00	25,219
# of Bedrooms	0.0786	-0.0000	0.00000	83.42	0.00	25,219
# of Livingrooms	0.0475	-0.0000	0.00000	50.44	0.00	25,219
FloorTotal	0.2046	-0.0000	0.00000	217.09	0.00	25,219

As Table 8 shows, the location of the building is associated even with building and apartment characteristics, therefore the spatial effects are not limited to neighborhood characteristics or distance variables. Therefore, the conceptual groups that the independent variables are put into in the next section should be interpreted accordingly that the non-locational variables could also have spatial dependence.

3.6.3 H1: Spatial dependence in residuals

A standard hedonic regression is employed first and hypotheses, starting with the existence of spatial dependence are tested with respect to the residuals of the regression. The regression model is defined in Table 9. The independent variables of the regression are conceptually grouped into 4: building characteristics, apartment characteristics, district variables and locational variables. For a more detailed explanation of variables, please refer to Appendix A.

Table 9. Regression Model 1

Regression Model 1: Base regression specification			
$$ln(p_i) = B_i + A_i + D_i + \varepsilon_i$$ p *denotes house prices* B *is a set of variables regarding structural characteristics of the building* A *is a set of variables regarding apartment characteristics* D *denotes the district dummy variables for fixed effects in the administrative areas*			
B	A	D	-
property_type	area	district	-
age	bathrooms		
heating_type	living_rooms		
construction_type	bedrooms		
gated	floor_at		

The residuals-fitted and Q-Q plots in Figure 17 show that the residuals of the regression exhibit some heteroscedasticity and deviate from normality by overestimating higher priced units. The heteroscedasticity and deviation from normality of residuals is neglected due to the nice asymptotic behavior of Moran’s I. The distribution of Moran’s I is found to be robust to violations in homoscedasticity and normality assumptions (Griffith, 2010).

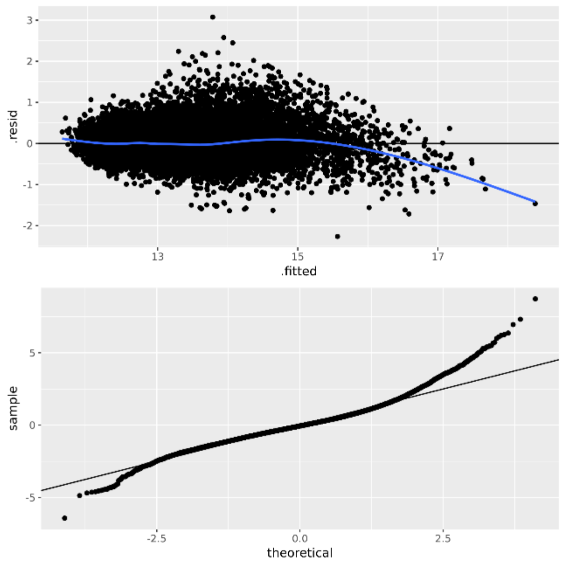

Figure 17 Residuals-Fitted and Q-Q plot for regression models Model 1

Table 10 shows the results of Moran’s I calculations employed to measure the spatial dependence in the residuals. The measure significantly differs from the expected value of the null hypothesis of no spatial autocorrelation given the variance, therefore the spatial dependence is inferred.

Judging from the variogram presented in Figure 18, the spatial dependence of the residuals happens within 3 kilometers of the units. The variogram for residuals differ from the variogram for prices in structure and size. The variogram for residuals does not exhibit structural instability after 30 kilometers. Its range of 3 kilometers is narrower than the range of 5 kilometers for prices. Its nugget and sill have a smaller value than the prices. These effects are in line with what is expected and confirms that the hedonic regression helps accounting for spatial effects.

Table 10. Moran's I of Regression Model 1 Residuals (Inference by normalization. Weight matrix is defined as unity standardization of inverse distances within 30 km bands.)

	Moran.I	Expected.MI	Variance.MI	Z.Score	P.Value	n
Model 1	0.1037	-0.0004	0.000001	114.6286	0	25, 219

The spatial dependence can be further analyzed locally. Figure 19 presents Local Moran's I measure grouped by building age and district. Newer buildings exhibit a higher spatial autocorrelation on average and mid-age buildings exhibit fewer clustering in unexplained prices. This might be due to the premium of younger buildings in Istanbul where a major earthquake is expected and old buildings that

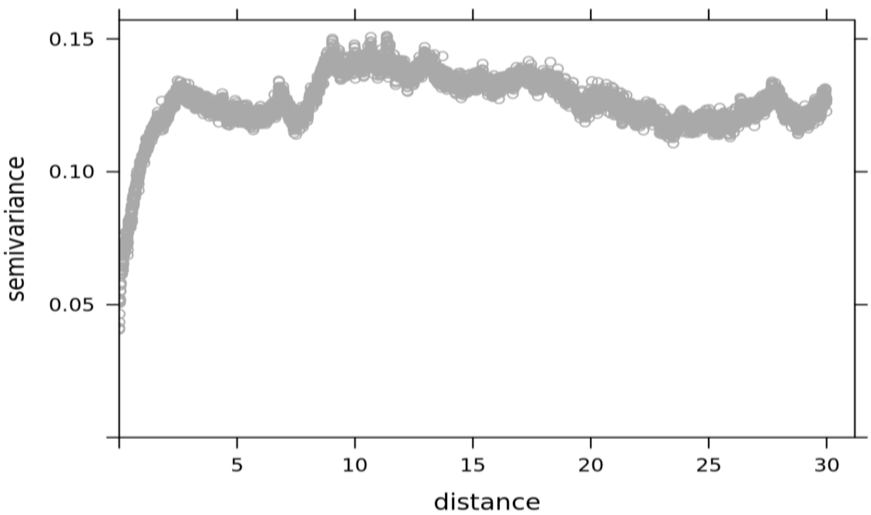

Figure 18 Variogram of regression residuals of Model 1, 5 meters sensitivity, sparse bins (n < 30) are deleted

have been constructed earlier than more recent regulations are deemed unsafe due to earlier construction practices.

Figure 19 shows that the spatial autocorrelation is different among districts as well. Two patterns are apparent in differences of unexplained spatial clustering of prices among districts. First, it seems that it is higher especially in areally large districts such as Eyüp and Sarıyer, where socioeconomically different neighborhoods are further away from each other and a binary district variable is not enough to capture fixed spatial effect. Second, it is higher in wealthier districts such as Beyoğlu, Beşiktaş and Kadıköy, where there resides highest priced luxury housing. In short, local spatial autocorrelation in districts are higher for districts where within district price differences are larger.

The robustness of results for H1 is checked using robust LM-based tests (Table 11) and results are found to be stable across the choice of spatial dependence and implied model. Note that Moran's I calculations above use unity-standardized spatial weight matrices and row-standardized spatial weight matrices are used in LM-based tests as these tests are developed using row-standardized spatial weight matrices. According to LM-based tests, there is significant spatial dependence in resisualds. It happens through both spatial error and spatial lag, implying a SARMA model as the spatial model.

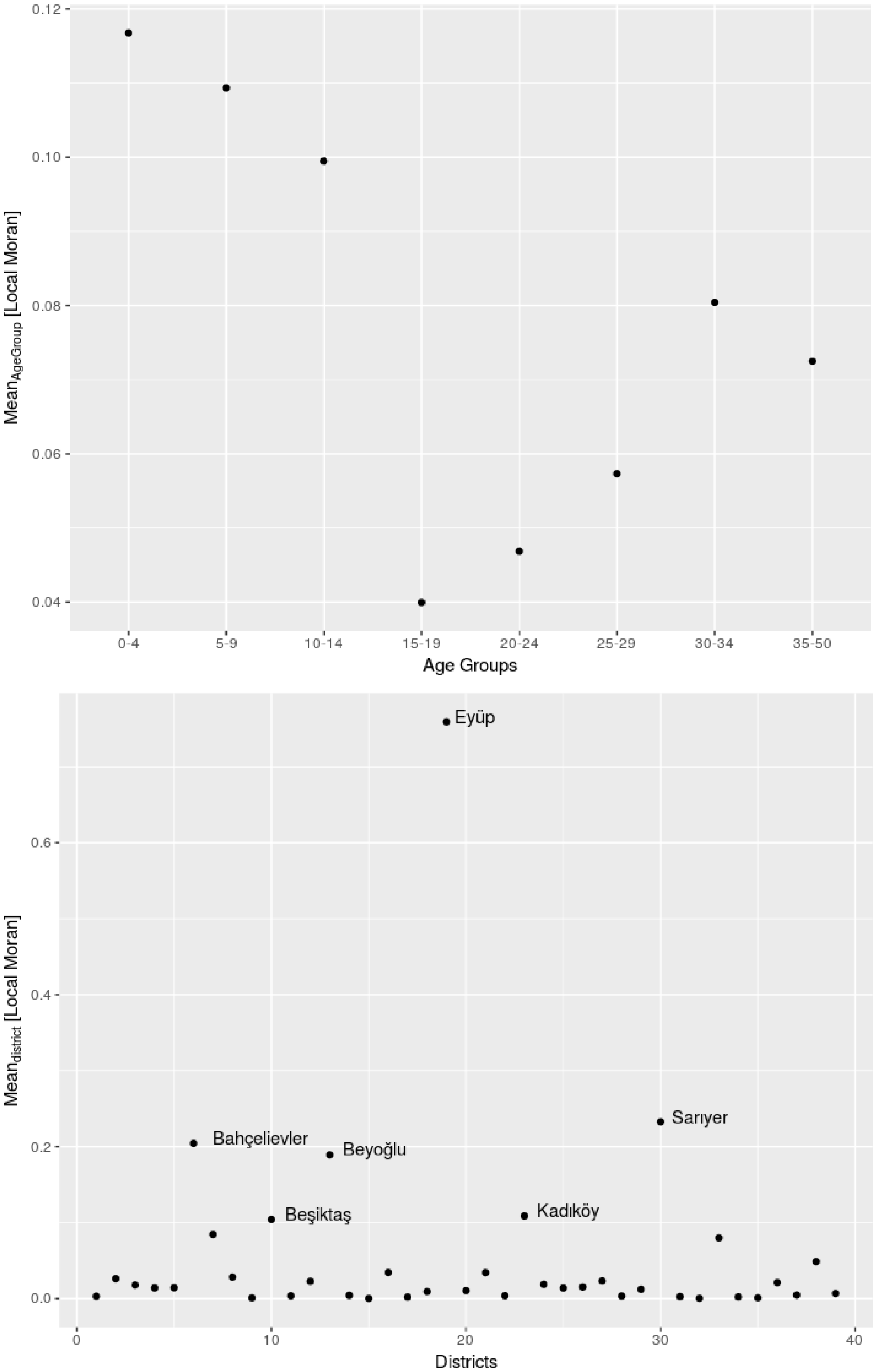

Figure 19 Average Local Moran's I in Model 1 residuals, grouped by building age (top) and district (bottom)

Table 11. Robust LM Tests for Spatial Autocorrelation in Regression Model 1 Residuals

	Robust LM-Error	Robust LM-Lag	Robust SARMA
Model 1	24,149	1,850	31,566
(p value)	0	0	0

3.6.4 H2: Construction process is a cause of spatial autocorrelation in residuals

Hypothesis 2 states that the construction process of the building is a cause of spatial dependence; spatially close buildings are similar in building age and share similar design features such as housing area and number of bedrooms. This hypothesis is plausible if there are some uncontrolled variables in the regression such as design or there are some non-linear effects of a controlled variable.

In order to check whether the premise that spatially close properties tend to be built around the same time, the relation between the building age and space can be examined. Table 8 in Section 3.8.2 presents this relation that there is a significant spatial autocorrelation of building age on space. Additionally, Table 8 also shows that other structural characteristics such as the number of living rooms, number of bedrooms, number of bathrooms and area are all spatially dependent. As these are positive autocorrelations, the findings support the premise that spatially close properties indeed tend to be built around the same time and that the structural features are spatially autocorrelated.

If the H2 were true, then we would have to observe a significant spatial relation between the residuals of the model and the structural characteristics. In order to test this, a bivariate Moran's I test is employed, details of which is explained in Section 3.5.1.2.3. Table 12 shows the result of bivariate Moran's I test between the residuals and the structural characteristics of age and area. Some possible characteristics such as the number of bedrooms, bathrooms and livingrooms are omitted in this analysis as they have low variation in themselves. Both age and area are found to have significant spatial relation with the residuals, which provides support for the hypothesis.

Table 12. Bivariate Moran's I between Regression Model 1 Residuals and Relevant Independent Variables (Inference by normalization. Weight matrix is defined as unity standardization of inverse distances within 30 km bands.)

	Moran.Bivariate.I	Expected.MI	Variance.MI	Z.Score	P.Value	n
Resid.-Age	0.0294	0	0.000000	62.0114	0	25, 219
Resid.-Area	0.0496	0	0.000000	104.6309	0	25, 219

In conclusion of H2, the findings in this analysis cannot reject the hypothesis that the construction process is a cause of spatial autocorrelation. The mechanism through which it happens seems to include a model mismatch, as the residuals are spatially dependent on variables that are already being controlled in the analysis.

The implications of H2 are of importance; the structural characteristics of a property is controlled in the standard hedonic regression equation given the assumption that the residuals are independent, though, they are spatially dependent on the control variables. This implies that (i) there might be omitted variables that are spatially correlated with the controls (ii) the standard hedonic model does not appropriately account for the impact of the spatial factors. Both implications are viable, however, further analysis of the results in Table 8 and Table C3 in Appendix C reveal that the latter is more evident. For example, spatial autocorrelation of age is stronger than that of area while the bivariate spatial dependence of residuals with age is weaker than that of area.

The robustness of the findings is checked via two additional methods following Chen (2015). In both methods, all variables and spatial weights are standardized and normalized. First, GSCI measure is calculated with respect to age, which is supposed to be similar to a Pearson-correlation coefficient thus indicating a positive spatial correlation. Calculated GSCI measure turns out to be similar to the calculated bivariate measure. Second, the local vector for the LSCI measure is calculated. In the case of no spatial relation, this vector is expected to be distributed around zero. The t-test rejects the hypothesis that the LSCI vector has a true mean of 0 with a t-statistic of 11.19, supporting findings of the previous tests that there is indeed a relation between the residuals and the independent variable spatially. Note that the relation is spatial but not in the original series. The LSCI mean is close to zero. The residuals have a mean of 0 and cannot statistically be differentiated than zero. Similarly, the standardized and normalized independent variable of age has a mean of 0 and cannot statistically be differentiated than 0. Moreover, both series do not show a standard correlation. Table 13 summarizes the findings of Chen's SCI indices. Even though residuals and standardized age are not statistically different from zero univariately and they lack bivariate correlation, their bivariate relation with respect to the space is different from zero.

Table 13. Bivariate Relations Between Regression Model 1 Residuals and Independent Variables, Including Chen's SCI indices

	Residuals-Age y = residuals of Model 1 z = std. and norm. Age	Residuals-Area y = residuals of Model 1 z = std. and norm. Area
GSCI (y,z)	0.0294	0.0496
LSCI (y,z) Sample mean	1.2e-06	2.0e-06
LSCI (y,z) t-statistic	11.19	25.76
y Sample Mean	5.0e-17	5.0e-17
y t-statistic	0.00	0.00
z Sample Mean	9.2e-18	2.3e-16
z t-statistic	0.00	0.00
Cor(y, z)	-8.0e-16	-2.0e-16
Cor(y, z) t-statistic	-0.00	-0.00

Second, in order to capture the size of the spatial effect, the residuals are tried to be explained spatially by a standard regression, following Chen (2015). The dependent variable is a spatial version of residuals. The independent variables are some of the structural characteristics of the apartment. Table 14 summarizes the findings. The structural characteristics included can significantly explain 2.3% of the variation in the residuals when a space definition is taken into account.

The inferences made in this section also holds for regression models 2, 3 and 4, albeit with smaller impacts. The results of the analyses for the other regression models are provided in Appendix B.

Table 14. Spatially Explaining the Residuals with Features of the Construction Process

	Dependent variable:
	nWresid
	Model 1
Age_norm	0.011***
	(0.001)
Area_norm	0.017***
	(0.002)
Bathrooms_norm	0.011***
	(0.001)
Livingrooms_norm	-0.001
	(0.001)
Bedrooms_norm	-0.008***
	(0.002)
FloorTotal_norm	-0.005***
	(0.001)
Observations	25,219
R^2	0.024
Adjusted R^2	0.023
Residual Std. Error	0.142 (df = 25213)
F Statistic	102.124*** (df = 6; 25213)
Note:	*p < 0.1; **p < 0.05; ***p < 0.01

3.6.5 H3: Proximity to social services is a cause of spatial autocorrelation in residuals

H3 states that social services such as the ones provided by fire departments, police stations, health institutions, and transportation services are shared among nearby houses. As they have an effect on house prices and not taken into account, they cause spatial autocorrelation.

This hypothesis is tested by including a new set of variables, L, to Regression Model 1, forming Regression Model 2. The new variables contain the distances between the social service provider and the location of the building defined by the specific geolocation of the unit. The new variables are shown under the group L in Table 15 and a full specification is provided in Appendix A.

Public transportation stops, shopping centers, police stations, fire stations, university campuses, and schools are considered as shared social services. Public transportation stops include 244 major stops such as metro stops, metrobus stops and airports across the city. Shopping centers include 149 malls across the city, which are culturally a common social environment especially for families with kids. Fire departments include 95 departments operated by the municipality and police departments include 218 departments operated by the government. Over 5400 schools are considered across the city in different variables depending on the education stage of the school and whether the school is public or private. 300 hospitals and 161 university campuses that are in the city are included in the analysis. Distances from the housing unit to the closest social service building is calculated for each social service and for each housing unit. Model 2 includes those set of distance variables, grouped as L, in addition to the variables that exist in Model 1.

8 of the 14 newly added variables are significant at the 0.05 level in the regression and they add an additional 1% to R2. The effects are such that primary school and universities have a positive impact on the price, where pre-schools and secondary schools have a negative impact. In addition to schools, only distances to fire stations are significant and the relation is positive. Regression coefficients and their standard deviations of the newly added variables are shown in Table 16. Full coefficients are given in Appendix B.

Table 15. Regression Model 2

Regression Model 2: Locational Variables Added to Base Specification			
$ln(p_i) = B_i + A_i + D_i + L_i + \varepsilon_i$ p *denotes house prices* B *is a set of variables regarding structural characteristics of the building* A *is a set of variables regarding apartment characteristics* D *denotes the district dummy variables for fixed effects in the administrative areas* L *is a set of variables based on specific geolocation of the building*			
B	A	D	L
property_type	area	district	Distances to closest major public transportation stop
age	bathrooms		Distances to closest shopping center
heating_type	living_rooms		Distances to closest police station
construction_type	bedrooms		Distances to closest fire station
gated	floor_at		Distances to closest university Distances to closest schools (of types defined by whether it is public or private and its education stage)

Table 16. Coefficients of Locational Variables in Regression Model 2 (Other Variables are Omitted in the Table for Brevity)

	Dependent variable: log(price)
	Model 2
...	
dist.2.closest.fire_department	0.00001^{***}
	(0.00000)
dist.2.closest.hospital	0.00000
	(0.00000)
dist.2.closest.shopping_mall	0.00001^{*}
	(0.00000)
dist.2.closest.police_department	−0.00000
	(0.00000)
dist.2.closest.transportation	0.00000
	(0.00000)
dist.2.closest.private.preschool	-0.00004^{***}
	(0.00001)
dist.2.closest.private.primschool	0.00002^{***}
	(0.00001)
dist.2.closest.private.highschool	-0.00004^{***}
	(0.00001)
dist.2.closest.private.middleschool	-0.00005^{***}
	(0.00001)
dist.2.closest.public.preschool	0.00000
	(0.00000)
dist.2.closest.public.primschool	0.00002^{*}
	(0.00001)
dist.2.closest.public.highschool	−0.00001
	(0.00001)
dist.2.closest.public.middleschool	-0.00003^{**}
	(0.00001)
dist.2.closest.university	0.0001^{***}
	(0.00000)
Observations	25,219
R^2	0.840
Adjusted R^2	0.839
Residual Std. Error	0.343 (df = 25136)
F Statistic	$1{,}607.943^{***}$
	(df = 82; 25136)
Note:	*p<0.1; **p<0.05; ***p<0.01

New variables do not improve the heteroscedasticity and the distribution of the residuals in a manner visible to the eye in residuals-fitted and Q-Q plots in Figure 20. The empirical variogram, on the other hand, exhibits small differences in both models, as seen in Figure 21. There is a slight decrease in range and sill in the variogram of Model 2, compared to the variogram of Model 1. Thus, visual inspection suggests that newly added locational variables decrease spatial autocorrelation in residuals, improving coefficient estimations; while having only a slight effect on fits and overall structure of residuals.

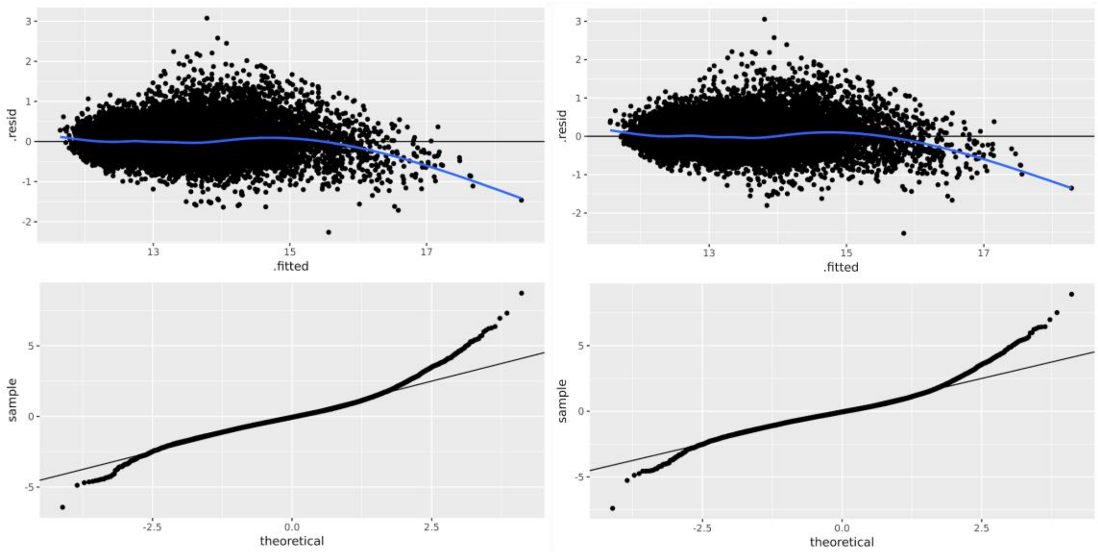

Figure 20 Residuals-Fitted and Q-Q plot for regression models Model 1 and Model 2, respectively

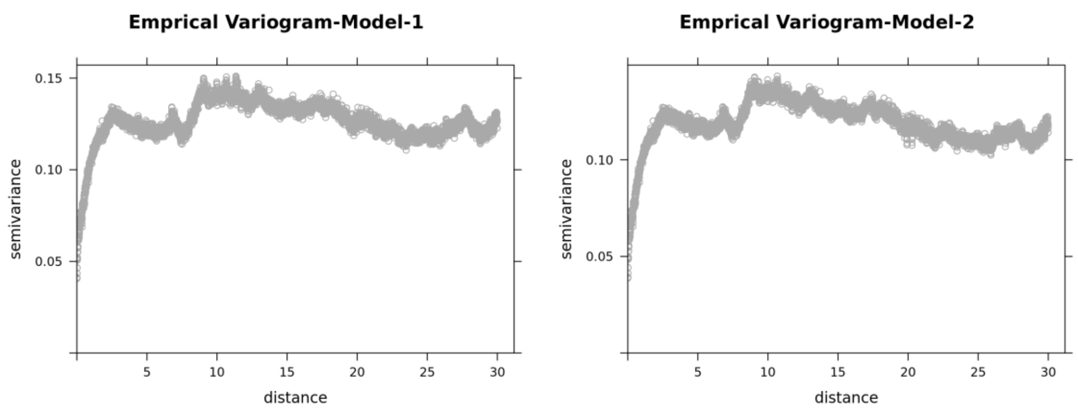

Figure 21 Variogram of regression residuals of Model 1 and Model 2 respectively, 5 meters sensitivity, sparse bins (n < 30) are deleted

Inference in using Moran's I is done by testing the null hypothesis of no spatial autocorrelation. Therefore, the inference only states whether there is autocorrelation present in the data. First two rows of Table 17 presents the results of Moran's I tests for Model 1 and Model 2 regressions. The last row of Table 17 shows the difference between Moran's I measures for both models and whether the difference is statistically significant given the variances. The details of the calculations are provided in the methodology section 3.3.1.1. Note that in judging the significance of the change in Moran's I statistics, the spatial weight matrices, W, are the same for two models and it's space interpretation is kept relatively intact by using unity normalization. Another point to consider is that even though Moran's I is not an estimator for rho of spatial error in general, it estimates rho well around zero and according to LM-based tests shown at Table 11, the data exhibits a very strong spatial error component in comparison to spatial lag component. As Moran's I is a measure of spatial dependence and the magnitude of the statistics is judged with respect to its moments, the data set used allows judging the magnitude of the statistic efficiently with small variances due to the fact that Moran's I measure converges well asymptotically.

Table 17. Moran's I in Regression Models Residuals (Inference by normalization. Weight matrix is defined as unity standardization of inverse distances within 30 km bands.)

	MI	Expected MI	Variance MI	Z Score	P Value	n
Model 1	0.1037	-0.0004	0.000000	114.63	0	25,219
Model 2	0.0691	-0.0005	0.000000	78.52	0	25,219
Δ(MI)	-0.0346			-38.14	0	

These findings are in line with those of Oust et al. (2019) which show that the addition of district level variables benefits the performance of the OLS significantly.

Moreover, their methodology of introducing artificial market districts with more homogeneous property pricing processes using k-means is similar to the inclusion of neighborhood level data in this study and yields similar results.

The results are checked using robust LM-based tests for both models (Table 18). The robust LM-based tests of spatial autocorrelation approve that there is strong spatial autocorrelation in both models. Note that Moran's I calculations in this study is performed with unity-standardized spatial weight matrices and LM-based tests are performed using row-standardized spatial weight matrices. According to LM-based tests, there is significant spatial dependence in resisualds in both models, albeit with a reduced magnitude when new locational variables are added.

Table 18. LM Tests for Spatial Autocorrelation in Regression Model 1 and Model 2 Residuals

	Robust LM-Error	Robust LM-Lag	SARMA
Model 1	24, 149	1, 850	31, 566
(p value)	0	0	0
Model 2	14, 643	1, 157	19, 158
(p value)	0	0	0

The inferences made in this section also holds for regression models 2, 3 and 4, albeit with slightly different impacts. The results of regression models are provided in Appendix C and the results of Moran's I for all regression models are provided in Appendix B.

3.6.6 H4: High-rise residential complexes are a cause of spatial autocorrelation

Sun, Tu, & Yu (2005) points out that there is a building dependence in addition to neighborhood dependence for multi-floor residential houses. All housing units in a multi-floor residence building have the same location, therefore they require an update in models based on neighborhood effects.

The total floor of a building is expected to be similar in spatially close buildings as the maximum height of a building is determined by city planners and the economic agents are expected to build buildings in heights that are most profitable to them. The spatial similarity of building heights is seen in spatial autocorrelation of the variable FloorTotal, which denotes the height of the building in floors. As given in Table 19, FloorTotal incorporates a significant level of spatial autocorrelation in itself and in fact it has the largest Moran's I statistic amongst all variables, including the price.

Table 19. Moran's I for FloorTotal Variable, Inference by Randomization

	Moran.I	Expected.MI	Variance.MI	Z.Score	P.Value	n
FloorTotal	0.2046	-0.00004	0.000001	217.0917	0	25, 219

If the number of floors is a cause of spatial autocorrelation, then we should observe a significant spatial effect of floorTotal variable on prices directly. For this purpose, Table 20 shows the spatial dependence of prices for two samples: all sample and low-floor buildings only, where low-floor buildings are defined to be buildings that are lower than 10 floors in height. The premise that spatial dependence is different based on whether the building is a high-rise building indeed holds in the data.

Table 20. Moran's I for Prices, All Observations and Low Floor Observations Only, Inference by Randomization

	Point MI	Expected MI	Variance MI	Z Score	P Value	n
All Sample	0.1315	-0.0000	0.00000	139.60	0.00	25,219
Low Floor Sample	0.1474	-0.0000	0.00000	102.20	0.00	19,204

The spatial autocorrelation of residuals of Model 1, for both the overall and low-rise buildings, are presented in the first two rows of Table 21. Low-rise subsample has a different Moran's I value in residuals as well than the overall sample. The null hypothesis of H4 is that height of the building has no impact; therefore, Moran's I for the high-rise (low-rise) sample equals that of the overall sample. To validate this baseline Moran's I with simulation, the analysis is replicated by performing the regression and computing respective Moran's I values for a random sample of 19,204 observations. This process is repeated a hundred times and an average Moran's I value is obtained, which is reported in the last row of Table 21. The average simulation Moran's I value is similar to that of the overall sample, however, the specific low-rise subsample has a different Moran's I value; rejecting the null hypothesis. Assuming the independence of residuals, the change in Moran's I is significant with respect to the variances. Note that the prices of apartments in low-floor buildings are more alike spatially than the prices of apartments in high-floor buildings, which suggest a negative spatial dependence effect of higher floor buildings.

Table 21. Moran's I of Regression Model 1 Residuals, All Observations and Low Floor Observations Only. Inference by Normalization. Third Row Indicates Average Results of 100 Random Samples

	MI	Expected MI	Variance MI	Z Score	P Value	n
MI residuals, all floors	0.1037	-0.0004	0.00000	114.60	0.00	25,219
MI residuals, low floors	0.1196	-0.0005	0.00000	86.23	0.00	19,204
MI sim, all floors	0.1040	-0.0005	0.00000	87.81	0.00	19,204

As opposed to other variables considered in the analysis section, the floorTotal variable is included in the regression models employed in an indirect way, that is through the variable floorAtClustered. floorAtClustered variable is a categorical variable that merges categorical floorAt variable and numerical floorTotal variable in a meaningful way. This transformation is performed due to the fact that the expected contribution of each additional floorAt to prices is not linear. An additional floor from the basement level to ground level has a high contribution to prices, where an additional floor from 15th to 16th level may matter marginally. Therefore, a floorAtClustered variable is formed using both floorTotal and floorAt variables, which is expected to meet common understandings of floors such as "underground level", "ground level', "first level", "in-between middle levels", "high level", and "top level". Thus, the floorTotal variable is indirectly included in calculations of floorAt variable and therefore cannot be included in the regression models by itself as doing so would cause multicollinearity among independent variables. The complete list of variables included in the regression models are provided in Appendix A.

The bivariate relation of the variable floorTotal and regression residuals verify the significant spatial dependence as well as the negative sign. As shown in

Table 22, Bivariate Moran's I between both residuals and floorTotal is -0.011, concluding that the high-floor buildings have negative spatial autocorrelation. Negative spatial autocorrelation means that residuals and building heights are particularly clustered in a dissimilar fashion across space. This is in line with the fact that high floor buildings are usually gated luxurious apartments, called residences, that are in general built on large lands on peripheral areas due the expensive nature of central land. As they add value by their luxury, community and living area, the link between the location and price disassociates, resulting in a negative spatial autocorrelation.

Table 22. Bivariate Moran's I of Regression Residuals and FloorTotal (Inference by Normalization. Weight Matrix is Defined as Unity Standardization of Inverse Distances within 30 km Bands.)

	Moran.Bivariate.I	Expected.MI	Variance.MI	Z.Score	P.Value	n
Resid.-FloorTotal	-0.0107	0.000000	-0.000000	-31.52	0	25, 219

The robustness of the results are checked using additional methods following Chen (2015) and reviewing the results under different regression models. Similar to those checks in previous sections, the bivariate relation between the Regression Model 1 residuals and floorTotal are examined using GSCI and LSCI indices. Results in Table 23 indicate that even though residuals and standardized floorTotal are not statistically different from zero univariately as well as they lack bivariate correlation, their bivariate relation with respect to the space is different than zero.

Table 23. Bivariate Relations Between Regression Model 1 Residuals and FloorTotal, Including Chen's SCI Indices

	Residuals-FloorTotal y = residuals of Model 1 z = std. and norm. FloorTotal
GSCI (y,z)	-0.0107
LSCI (y,z) Sample mean	-0.0000
LSCI (y,z) t-statistic	-3.7788
y Sample Mean	0.0000
y t-statistic	0.00
z Sample Mean	-0.0000
z t-statistic	-0.00
Cor(y, z)	-0.0086
Cor(y, z) t-statistic	-1.36

A regression on the residuals verify the analysis. FloorTotal variable alone can significantly explain 0.1% of the variation in the residuals spatially and the coefficient is negative (Table 24). However, the analysis is only true for regression models 1 and 2. In models 3 and 4, which include neighborhood fixed effects in addition to districts, the relation changes sign and the impact decreases. The change of effect in models 3 and 4 can also be seen in the bivariate Moran's I in Table 25. Even though the bivariate spatial relation is significantly negative for model 1 and 2, it is significantly positive for model 3 and 4.

The results indicate that while the negative spatial effect is present within districts, it is accounted for in neighborhood fixed effects. Neighborhoods are the smallest administrative areas in the city and each district has more than 10 neighborhoods in the city on average as there are 39 districts and 517 neighborhoods. It might be the case that the high-luxury residences are causing negative spatial autocorrelation as externalities within the districts while outnumbering the apartments on the market within the neighborhoods.

Table 24. Spatially Explaining the Residuals: FloorTotal. This Regression Shows that FloorTotal can Explain 0.1% of the Variation in the Residuals when a Space Definition is Taken into Account

	Dependent variable:			
	nWresid			
	Model 1	Model 2	Model 3	Model 4
FloorTotal_norm	-0.004***	-0.004***	0.001***	0.001***
	(0.001)	(0.001)	(0.0005)	(0.0003)
Observations	25,219	25,219	25,219	25,219
R^2	0.001	0.002	0.003	0.001
Adjusted R^2	0.001	0.002	0.0003	0.0005
Residual Std. Error	(df = 25218) 0.144	0.094	0.073	0.055
F Statistic	(df = 1; 25218) 17.409***	56.069***	7.498***	12.999***
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table 25. Bivariate Moran's I of Regression Residuals and FloorTotal. (Inference by Normalization. Weight Matrix is Defined as Unity Standardization of Inverse Distances within 30 km Bands.)

	Bivariate MI	Expected MI	Variance MI	Z Score	P Value	n
Resid.-FloorTotal						
in model 1	-0.0107	0.0000	0.00000	-31.52084	0	25,219
in model 2	-0.0130	0.0000	0.00000	-42.82406	0	25,219
in model 3	0.0046	0.0000	0.00000	5.095859	0	25,219
in model 4	0.0046	0.0000	0.00000	5.204712	0	25,219

3.7 Conclusion

The standard hedonic regression of real estate prices has a high R^2 (0.83), however, this is misleading for multiple reasons. First, the residuals are heteroskedastic even with a logarithmic transformation of the dependent variable. Second, there is spatial autocorrelation observed in regression residuals, which inflates the explanatory power as measured by R^2. The coefficients of the covariates in the regression are overstated for the same reason. Moreover, a high number of areal dummies contribute largely to the explanatory power of the regression, which undermines parsimony and the need for a regression.

An analysis on the variables show that housing prices, as well as the characteristics of a property including building age, area and number of bedrooms are highly spatially autocorrelated. In other words, given a spatial definition amongst observations, many variables including price and perhaps less intuitively independent variables such as age or area are autocorrelated.

The spatial autocorrelation tests show significant spatial autocorrelation in residuals of pricing models as well and the results are robust to using different measures. Local constituents of this spatial autocorrelation have uneven breakdowns in variables such as age or district. Empirical variogram can be plotted meaningfully with 5 meters precision as the data includes the geolocation of observations and the variogram shows that the spatial autocorrelation takes effect within a few kilometers around the housing. Based on the empirical variogram plots, the range of the spatial dependence is around 5 kilometers for prices, 2 kilometers for residuals of regression models controlling for district fixed effects, and 1 kilometer for residuals of regression models controlling for neighborhood fixed effects. The results indicated by measures and the empirical variogram plots seem consistent. Depending on the regression models used, 1% to 2.5% of the variation in residuals could be explained spatially by the independent variables.

Results coming from a number of spatial autocorrelation tests seem to indicate spatial dependence in both spatial error and spatially lagged terms, implying a SARMA model for spatial regression. Based on the robust LM-based tests, the spatial error component is much larger than the spatial lag component, implying that the effect of unobserved or misaccounted characteristics are higher than the effect of reaction of prices with each other.

Several hypotheses seen in literature regarding the causes and mechanism of spatial dependence is tested. According to the results in which multiple techniques are used including regressions for particular subsamples, spatial bivariate relations between residuals with covariates or spatial explanation of residuals by covariates; the data cannot reject these hypotheses and provides support for them.

Specifically, the findings indicate that one cause of the spatial dependence is the construction process of the building. Results show that not only spatially proximate buildings tend to be of similar age but also the design features of the property tends to be clustered in space as building and apartment characteristics are highly spatially autocorrelated. Notably, if the construction process hypothesis holds, then there exists spatial cross-correlation between age and regression residuals. The empirical results show that there indeed exists spatial cross-correlation between age and regression residuals as age can spatially explain 0.5% of the variation in regression residuals. Furthermore, results point to that there is a non-linear effect of age as seen in the spatial autocorrelation in different age subsamples, indicating that the impact of age of a housing unit is not fully accounted for in a standard hedonic regression.

Another cause for spatial dependence is shared social services that are not generally incorporated in models. When geolocational proximities to shared social services such as public transportation, shopping centers, police stations, fire stations, schools and universities are included in the analysis, spatial autocorrelation significantly reduces. The significant reduction in spatial dependence in addition to a smaller increase in the explanatory power of the model suggests that missing social services data in hedonic models, particularly those related to schooling in the data of this study, can contribute to spatial dependence and be an alternative for including locational fixed-effect variables.

The high-rise residential complexes are also a cause of spatial dependence. The height of the buildings are spatially autocorrelated, even more so than the price of the units. Additionally, low-rise buildings have higher spatial clustering in prices, suggesting that the high-rise residential complexes in general represent price oddities in their environment. Regression residuals and the total number of floors are spatially correlated, providing further support. The height of the buildings can spatially explain 0.1% of the variation in regression. Furthermore, the findings suggest that the high-rise residential complexes are at oddity with the housing units in their district but they are the dominant housing units in their neighborhoods.

CHAPTER 4

EVEN PRICING PHENOMENA IN REAL ESTATE

4.1 Abstract

Cross-sectional house price distributions are expected to be leptokurtic. However, fat tails are not the only inconsistency to the normal distribution assumption in housing prices. Empirically, the distribution is heavily clustered at particular points which results in spikes in a histogram. A further examination of the prices also reveals that the use of detailed figures in pricing is uncommon, the prices are clustered around more even figures.

Since the real estate market is an important part of the economy and the real estate purchasing decision usually involves an important share of the wealth of the buyer, the existence of the tendency towards price clustering, its nature and economic impact is worthy of analysis. This study documents and formally shows the pricing anomaly in real estate markets in terms of the tendency to anchor prices to even numbers. Using data from multiple markets, it is claimed that cognitive ease, negotiation and culture could explain the pricing behavior, while price-levels, information friction and taxes cannot fully explain it. To our knowledge, this is the first study to investigate even pricing as a pricing anomaly in the real estate market using a unique data set and for the Istanbul property market.

4.2 Introduction

Cross-sectional house price distributions are expected to be leptokurtic (Ohnishi et al., 2011). However, fat tails are not the only inconsistency to the normal distribution assumption in housing prices. Empirically, the distribution is heavily clustered at particular points which results in spikes in a histogram. Figure 22 demonstrates the different ways of putting the same set of prices in the data of this study to a histogram. When the bin width is increased, spikes at particular points are observed.

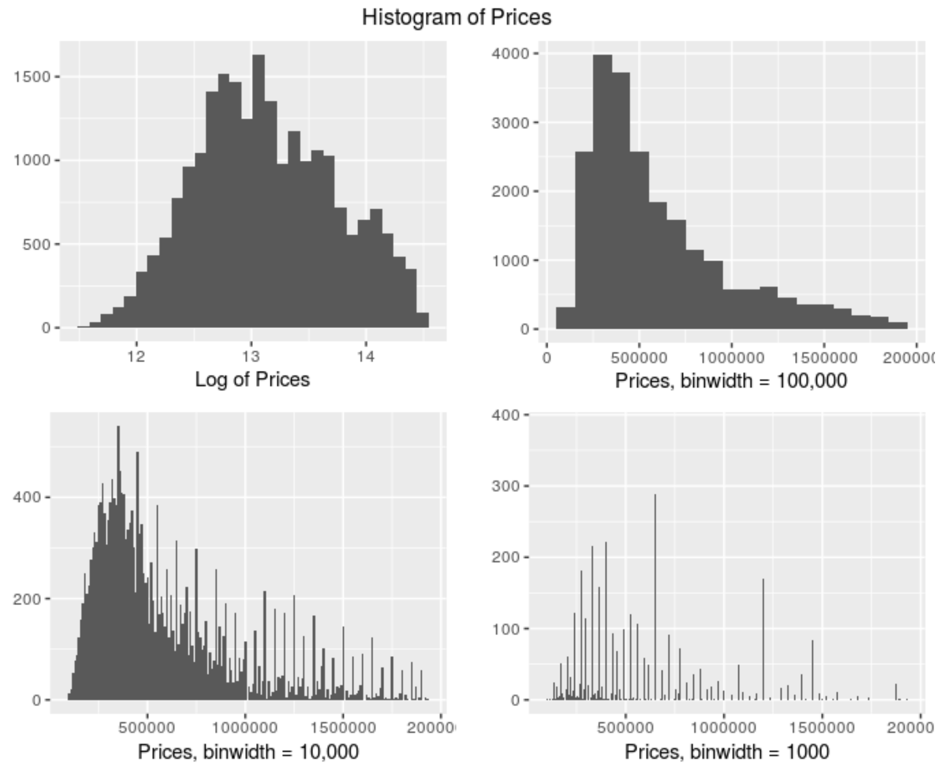

Figure 22 Different histogram plots of the same data

A further examination of the prices reveal that not only is the use of detailed figures in pricing such as $313,875 uncommon but also the prices are clustered around more even figures such as $320,000 rather than $318,000. This type of clustering in prices ending with zeros is called even, even-ending or round pricing in the literature whereas prices below even-ending are called just-below or charm pricing.

As detailed in Section 1.3, the real estate markets are an important part of the economy. It is bigger in size than the well-studied equity markets, the market participation rates are higher in real estate and the real estate activities accounts for a significant portion of GDP. In addition to the depth and penetration of real estate markets, the purchasing of a real estate is one of the most important decisions for most individuals as it includes a significant portion of the individual's wealth. Thus, the existence of the tendency towards price clustering, its nature and economic impact is worthy of analysis. In particular, the effect of even pricing would be around a few percent of the house price, though, even one percent of the median house prices is an economically significant amount which is comparable to months' worth of disposable median income. For instance, in the data set of this study, the median asking price of a multi-unit housing in Istanbul as of October 2017 is 490,000TRY. One percent of the median house price is 4,900TRY where the monthly disposable median income in Turkey is 1,324TRY.[20] Consequently, any inefficiencies in the market is of direct importance to the general well-being.

The research question of this chapter is whether there is a pricing anomaly in the real estate market in terms of the tendency to anchor prices to even numbers. In order to study the real estate prices, first, price clustering patterns are documented, then the pricing equation for the residential property market is formed and predicted values are obtained from the model so that these values can be compared to the real

[20] Turkstat annual equivalised household disposable income statistics of 2017. (http://www.turkstat.gov.tr/PreIstatistikTablo.do?istab_id=1602)

prices for different groups of price endings. To be able to perform these analyses, the observations are grouped based on their price-ending patterns and any discussion will be based on these groups of observations.

4.3 Literature review

Price anomalies are a central part of contemporary finance research. Price anomalies are an important part of finance research. The developments in psychology, behavioral economics and finance has emerged over the last few decades, contributing to the discussion on price anomalies. One of the important works in the real estate literature studying these anomalies is by Palmon et al. (2004), which documents and analyzes clustering in real estate prices. The authors group house prices according to the last three digits of the list prices. In their sample, more than half of the observations are “just-below” prices and 20% of the houses have even-ending pricing with three zeros. Beracha and Seiler (2014) report similar figures in their sample. They document that more than 75% percent of the prices are associated with a round or “just-below” pricing.

One theoretical explanation to price clustering is the cost of information acquisition. Investors use a coarse pricing set due to informational equilibrium, where the marginal cost of acquiring additional information equals the marginal benefit of using it. Using a coarse pricing set results in a price clustering (Ball et al., 1985). In other words, prices are conjectured to deviate from their fundamental values due missing or costly information.

Another possible explanation to price clustering regards the cognitive ease of using round numbers. People naturally rely on round numbers, such as multiples of 10, as cognitive reference points (Rosch, 1975). In the finance context, Kuo et al (2015) treats the reliance on even-ending orders as a cognitive limitation. They find that cognitive limitation of an individual investor, measured by the frequency of his/her limit order submissions for the stock market that are ending with 0, is correlated with investor's performance. Lacetera et al. (2012) show that paying attention to even numbers can take effect in the used-car market and report significant drops in transaction prices due to even mileage thresholds. They attribute the effect to the inattention of buyers.

Literature suggests that negotiation processes can also be of importance for price clustering. Harris (1991) studies the stock markets and theorizes that if traders use discrete sets of pricing in order to simplify their negotiation, price clustering exists. On the contrary, findings by Cardella and Seiler (2016) based on an experimental setting in the real estate field show that the use of rounded figures seem to slightly increase the total number of offers needed to reach an agreement and increase the length of the negotiation process.

One possible explanation to the use of just-below even-ending prices is the affordability phenomena, that is, the price of the property is perceived to be less than it already is when the price is tagged just-below even prices. In fact, Palmon et al. (2004) states that the tendency to use just-below and even-ending prices is positively related to the number of listings a broker has. This is in line with the understanding in retail marketing, where just-below pricing is widely used. However, Nguyen et al. (2007) argues that in retail, price ending practices are not universal and cultural context could affect perceptions.

4.4 Data and methodology

The dataset used in this study consists of over 300,000 observations of real estate listings across Istanbul as of October 15, 2017. The data is from HurriyetEmlak.com. HurriyetEmlak.com is a well-known marketplace for real estate listings in Turkey as it is one of the two leading online portals and has a large coverage of the market across Turkey. The data set includes all listings in Istanbul that the provider had at the cross-section; therefore, it is considered to be robust to sampling bias and inflationary time effects.

Each observation consists of an attribute set over 20 variables including features like asking price, the area in square meters, property types, date added, building age, number of floors in the building, number of bedrooms and bathrooms, and the total number of rooms. The data includes multiple real estate markets such as residential, commercial and land markets for sale and for rent. Main housing property types are flats, luxury residences and single houses. Main commercial property types are stores, common offices and luxury offices. The location of an observation is given by the hierarchical administrative areas that the property belongs to in terms of the neighborhood and by the geolocation of the property specifically. Figure 23 shows a mapping of the sale listings for housing in Istanbul.

A significant characteristic of the data set is that the prices are listing (asking) prices. Transaction prices, to my knowledge, are not readily available for the Turkish real estate market to validate the analysis. Governmental records are neither public nor fully accurate due market practices aiming tax evasion, bank records are based on expert valuation rather than transaction data and the data sets of the real estate brokerage agencies is at a high risk of non-random sampling.

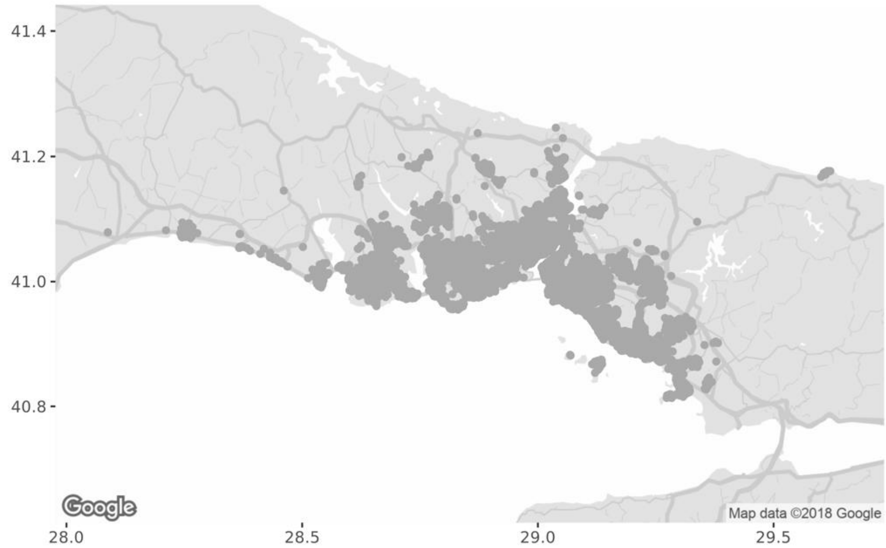

Figure 23 Locations of the sale listings for housing in Istanbul

Nevertheless, the even pricing phenomena in asking prices are still of interest even in the case that transaction prices do not exhibit the same phenomena since it may have impact on the market dynamics as explained by studies such as Cardella and Seiler (2016). Second, to our knowledge, there is no evidence that the transaction prices would have different characteristics than asking prices in terms of their evenness considering the findings in the literature regardless of their respective behavioral orientation for explanation. The rational theory shows that informational frictions and costs cause the usage of coarse information sets. Similarly, the behavioral theory explains that cognitive ease causes the use of even numbers. Both theories predict a similar phenomena to be observed in prices even after the negotiation process. Moreover, it is plausible that the negotiation process might amplify the tendency to use even pricing, as the negotiation process takes up significant cognitive load for both buyer and the seller, and cognitive ease is one reason for using even prices. In fact, findings of Palmon et al. (2004) show that the frequency of price ending of '000' is 19.55% in asking prices whereas it is as high as 50.49% in transaction prices.

The listing prices could differ from the transaction prices due to the margin for negotiation. If the margin for negotiation is distributed idiosyncratically across the real estate market then, on average, we can deduce the transaction price. However, if the average margin for negotiation differs, for example, by region, then this difference may cause a bias on the observed prices in the respective regions. One reason why the average margin for negotiation can be different is the liquidity of the property. Liquidity is measured by the time on the market (TOM) (Jud et al., 1996). The literature suggests that atypical properties and properties with higher asking prices relative to their transaction prices tend to have higher TOM hence lover liquidity (Jud et al., 1996; Knight, 2002; Merlo & Ortalo-Magné, 2004; Yavas & Yang, 1995). Consequently, in an effort to alleviate the problem, the listings that are older than 2 months are excluded from the analysis which helps us to downweigh the effect of liquidity and reduce the possible differences in the margin for negotiation.[21]

In analysis of the effects under question, the grouping of observations is done as follows. Let the data set S to refer to the collection of our listing observations:

$$S = \{Listing_i\}, i = \{1,2,3\ldots, number\ of\ observations\} \quad (56)$$

We define a model of price-endings for listings. For instance, for a model of prices endings where the last 4 digits are all zeros would be:

[21] The cut off point of 2 months is due to the longest time allowed for a listing to be active on HürriyetEmlak.com without a renewal action.

$$m_k = \{ Xk0000 \mid X \in \mathbb{N}_1 \}, k = \{0,1,2\ldots9\} \quad (57)$$

Then, if we group the observations based on the last fifth digit of the price of the specific listing, P_i , we have S_k as:

$$S_k = \{Listing_i \mid P_i \in m_k\}, k = \{0,1,2\ldots9\} \quad (58)$$

With this procedure, we separated D into eleven mutually exclusive and collectively exhaustive subsets, where the eleventh set being the observations that do not have all zeros at the end of their price tags.

$$S = \{S_k, S_{other}\}, k = \{0,1,2\ldots9\} \quad (59)$$

The price ending model in Equation 57 is one of the several price-ending models used in the analysis section. Given the groupings, the analyses are based on inter-group differences and standard ANOVA methods are used.

For base analysis, the dataset is divided into two: training and prediction samples, which respectively contain 15% and 85% of the housing sales sample with 6 digit prices. Coefficients are estimated using the training sample which constitutes and rest of the analysis are performed using the prediction sample. The pricing model is a standard hedonic regression model and expressed in Equation 60. The log of prices are regressed on a set of independent variables covering the structural, locational, neighborhood and interior features of the housing. Full model definition and results are given in Appendix D.

$$log(P) = \beta_{training}X + u \quad (60)$$

When using the prediction data set with the resulting coefficients of the regression run on training data set, the goodness of fit calculation slightly changes. The residuals and the R^2 in the prediction set are calculated as detailed in Equations 61-64.

$$u_{prediction} = log(P_{prediction}) - \hat{P}_{training} \quad (61)$$

$$SSR = \Sigma {u^2}_{prediction} \quad (62)$$

$$SST = \Sigma(log(P_{prediction}) - log(\underline{P})_{training})^2 \quad (63)$$

$$R^2{}_{prediction} = 1 - SSR / SST \quad (64)$$

4.5 Analysis

4.5.1 Base analysis

We examine empirically whether there is a pricing anomaly in the real estate market in terms of the tendency to anchor prices to even numbers. Different real estate markets are investigated separately and price-ending patterns of the observations are analyzed. In order to carry out the analysis, the number of digits of the prices for each analysis should be the same. For this reason, first, the observations with 6-digit prices in the housing market for sale is focused on as a base analysis.

Median sales price for housing is 490,000TRY in the dataset. It is a 6-digit figure and close to the middle of all 6 digit numbers. The mean sales price is

682,741TRY and it is also a 6-digit figure. In fact, 81.6% (18,761) of all observations in the data have 6-digit figures.

When looked at the distribution of the price endings of this subsample defined by the 6-digit prices in housing sales, it is observed that the prices are distributed with an increment of 5,000. This is seen in Figure 24, where a bar plot of the counts of observations grouped by their last 5 digits of the prices is given. Price endings outside this increment of 5,000 constitutes only a marginal amount of the observations, as no other group alone form up to 1% of all observations.

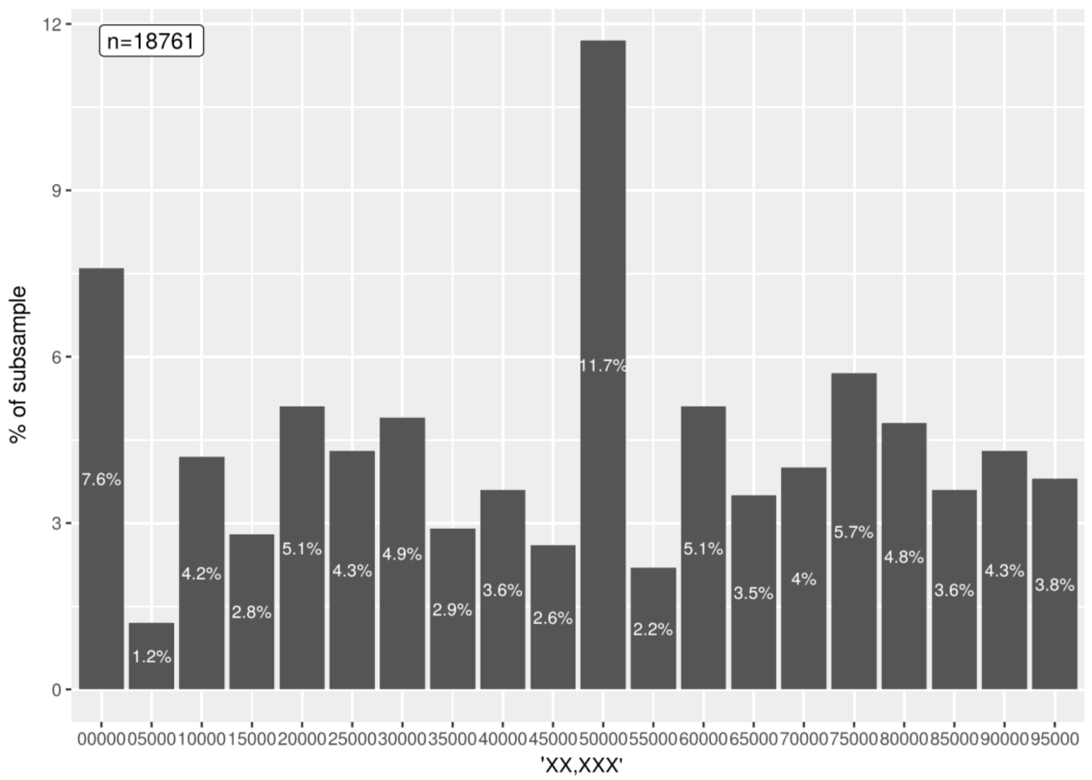

Figure 24 Distribution of price ending for housing sales subsample, model: dXX,XXX, bins < 1% excluded in the plot

Another interesting point in Figure 24 is that some price endings, particularly even price endings, are more popular than the others. There are more observations with ‘00,000’ or ‘50,000’ price endings at the expense of the price-ending groups around these even figures; supporting the view that people are rounding the price figures around the even numbers even at high stakes.

In analyzing this subsample, we focus on the most even figures, prices ending with zeros, and how pricing behavior around these figures change. The observations are grouped using the procedure explained in Equations 56-59. First, they are grouped into bins depending on whether their last 4 digits are all zeros. If their last 4 digits are ‘0,000’, then they are grouped into bins defined by their 5th digit from right. If their last 4 digits are not all zeros, they are put into the “other” bin.

Perhaps surprisingly, this grouping puts most of the observations into meaningful bins defined by the last 5th digit. For instance, it does so for half of the groups in Figure 24 for endings such as ‘00,000’, ‘10,000’, or ‘50,000’; and it sorts out the remaining groups such as “05,000’, ‘25,000’, or ‘75,000’ into “other” bin. Overall, 55.3% of the observations in 6-digit subsample end with ‘0,000’, as shown in Figure 25, and put into respective bins defined by the 5th digit, where the remaining observations are put into the “other” bin.

Focusing on the behavior of pricing around the even price endings, which is the 55.3% of all observations with 6-digits, the bins are sized as shown in Figure 26. Figure 26 is significant because it nicely documents even pricing at an economically significant level. The observations in the subsample are grouped based on their second largest digit, where a unit change in this digit economically means 10,000TRY, which corresponds to 1 to 10 percent of the house price depending on the house price level. 10,000TRY is 2% of the median house price and 63% of the disposable yearly income in Turkey.

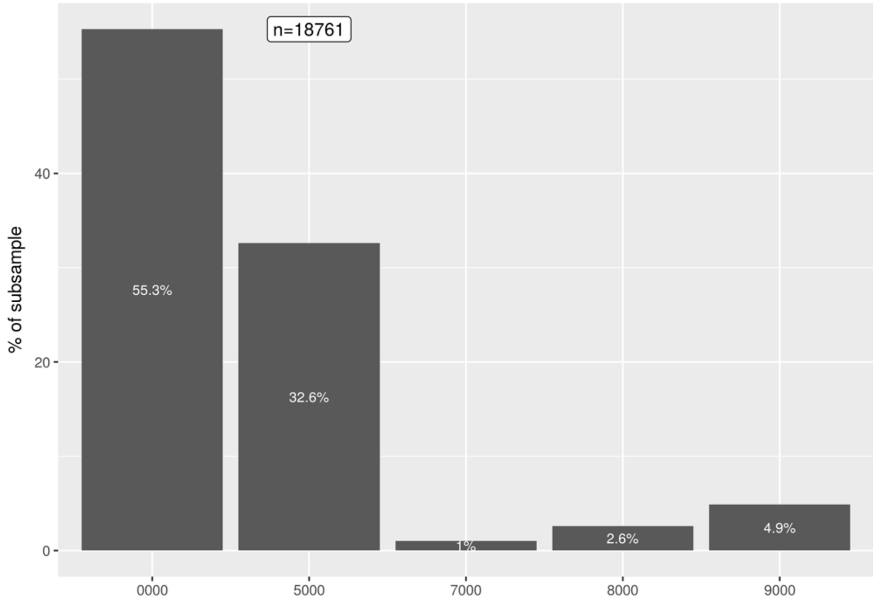

Figure 26 Distribution of price ending for housing sales subsample, model: ddX,XXX, bins < 1% excluded in the plot

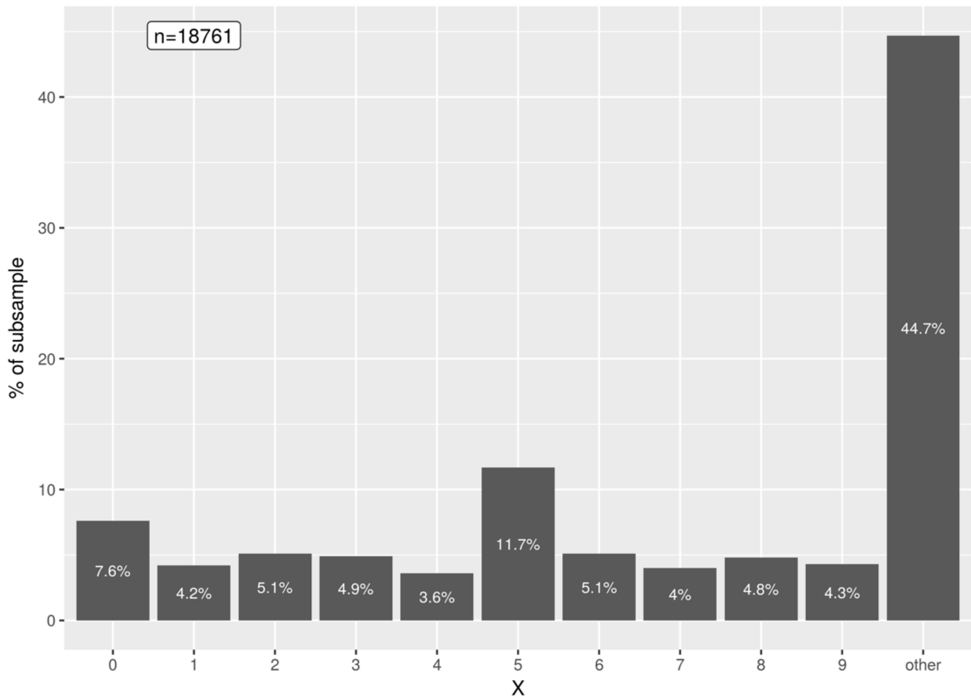

Figure 25 Distribution of price ending for housing sales subsample, model: dX0,000

Even amongst the observations with price endings of ‘0,000’, the fifth digit clusters around 0 and 5. If the clustering is examined leaving the ‘other’ bin out, 24.7% of the observations whose prices are 6-digit numbers and have all zeros in the last four digits have a 0 in the fifth digit. Similarly, 31.2% of such observations have 5 in the fifth digit. The frequency of the rest of the digits on the fifth digit is less than what a uniform distribution would predict, that is less than 10%, excluding the “other” bin.

These statistics not only confirm the clustering observed in the initial histogram that were present as spikes but also explain the patterns in pricing behavior. This pricing behavior may also have an economic meaning. Particularly, when all 4 digits of 6 digit prices are zeros, the number of observations with the 5th digit equals to 0 or 5 are more than the others. In other words, there is an increment of ‘50,000’, in addition to the increment to ‘5,000’. It is here documented that there are major clustering patterns at the second most important digit of the prices as well.

To judge the impact of the clustering, a pricing model is formed (Equation 60). The model is first applied to the smaller training set, where the R^2 of the fit is 0.75. Then the coefficients are used in the larger prediction set, where the R^2 of prediction sample is 0.72 (Equations 61-64) and the regression residuals display a slight heteroscedasticity (Figure 27) that is characteristic in standard hedonic regressions in real estate.

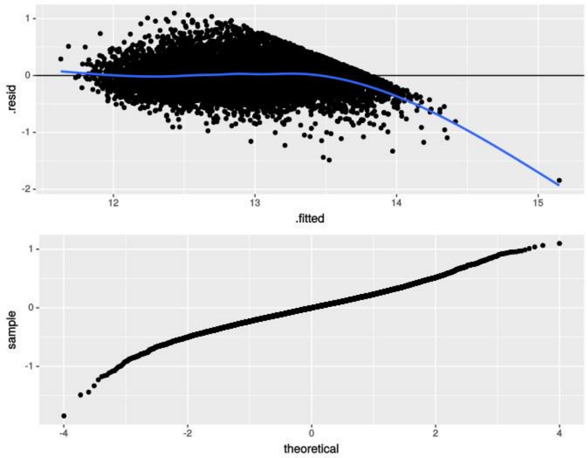

Figure 27 Residual-fitted and Q-Q plot for predictions (Housing sales sample, 6 digit price tags only)

]The residuals of this regression is grouped based on the observations' price ending patterns, following the same procedure explained in Equations 56-59. The residuals are then regressed on the categorical group variable, D, as in Equation 65.

$$u_{prediction} = \beta_0 + \beta_1 D + v \qquad (65)$$

The residuals of this regression are compared across groups. Figure 28 lists the means and standard deviations of the residuals within the groups and a box plot for a visual inspection.

Last 5th digit	Last 4 digit	n	Residual mean	Residual SD
0	0000	1210	0.0676	0.2641
1	0000	666	-0.0318	0.2382
2	0000	812	0.0126	0.2554
3	0000	779	0.0039	0.2260
4	0000	572	-0.0127	0.2439
5	0000	1873	0.0717	0.2546
6	0000	813	0.0053	0.2499
7	0000	643	0.0059	0.2578
8	0000	772	0.0250	0.2317
9	0000	684	0.0299	0.2533
other	XXXX	6849	-0.0357	0.2361

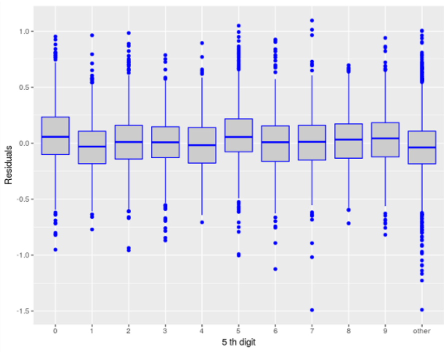

Figure 28 Pricing equation residuals by price-ending groups

As the 5 digit is not supposed to be a determinant of prices, the residual means of the groups should be the same and equal to zero. An ANOVA analysis is employed to test whether the means of the residuals are jointly zero. The test, for which the results are displayed in Table 26, rejects the null hypothesis with a p-value of 0.0000 that the group means are jointly zero.

Table 26. Analysis of Variance of Pricing Equation Residuals by Price-Ending Groups

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Df	2	7,836.000	11,067.640	10	3,923	11,749	15,662
Sum Sq	2	485.280	649.857	25.762	255.521	715.039	944.798
Mean Sq	2	1.318	1.779	0.060	0.689	1.947	2.576
F value	1	42.705		42.705	42.705	42.705	42.705
Pr(>F)	1	0.000		0.000	0.000	0.000	0.000

Similarly, the residuals from the pricing equation is regressed on the categorical digit group variable D. Table 27 shows the results of the regression which supports the findings of the ANOVA test. The group variable D is significant in predicting the residuals and explains about 2.7% of the overall variation in residuals. The coefficients on the digit group variables of zero and five have the highest impact and are significant at the 1% level.

4.5.2 Analyzing prices with non-’0000’ ending

The base analysis looks at the second most important digit in the prices where the price ends with ‘0000’. The observed phenomena comes from the majority (55.3%) of the observations that are put into bins 0 to 9, and the remaining observations, being put into the 'other' bin, did not contribute positively to the observed effects. However, the 'other' bin is covered here for a complete analysis and it is found that there exists similar phenomena within this 'other' bin as well.

Table 27. Explaining Pricing Regression Residuals by Clustering Around Even Prices

	Dependent variable:
	residuals
	OLS
digit_group0	0.106***
	(0.008)
digit_group1	0.003
	(0.010)
digit_group2	0.047***
	(0.009)
digit_group3	0.039***
	(0.009)
digit_group4	0.025**
	(0.011)
digit_group5	0.114***
	(0.006)
digit_group6	0.045***
	(0.009)
digit_group7	0.045***
	(0.010)
digit_group8	0.056***
	(0.009)
digit_group9	0.055***
	(0.010)
Constant	−0.036***
	(0.003)
Observations	15,673
R^2	0.029
Adjusted R^2	0.028
Residual Std. Error	0.244 (df = 15662)
F Statistic	46.706*** (df = 10; 15662)
Note:	*p<0.1; **p<0.05; ***p<0.01

As shown in Figure 25 in the previous section, a large majority of the observations that do not end with all zeros and thus put into the ‘other’ bin have a price ending of ‘5,000’. When looked into detail, these which by themselves inhabit clustering of the preceding digit at 2 and 7. The distribution of the preceding digit in this sample is given in Figure 29. Clustering around ‘25,000’ and ‘75,000’ are other forms of round pricing; however, what is more interesting is that this time the clustering occurs at the expense of the digits 0 and 5, in contrast to even pricing around all zero’s shown above which is at the expense of digits next to 0 and 5 (see Figure 26). This phenomena of 0 being distinctively unpopular as a preceding digit in a higher level of precision suggests that there is an avoidance of being too precise around round figures in pricing in addition to the tendency to use round figures.

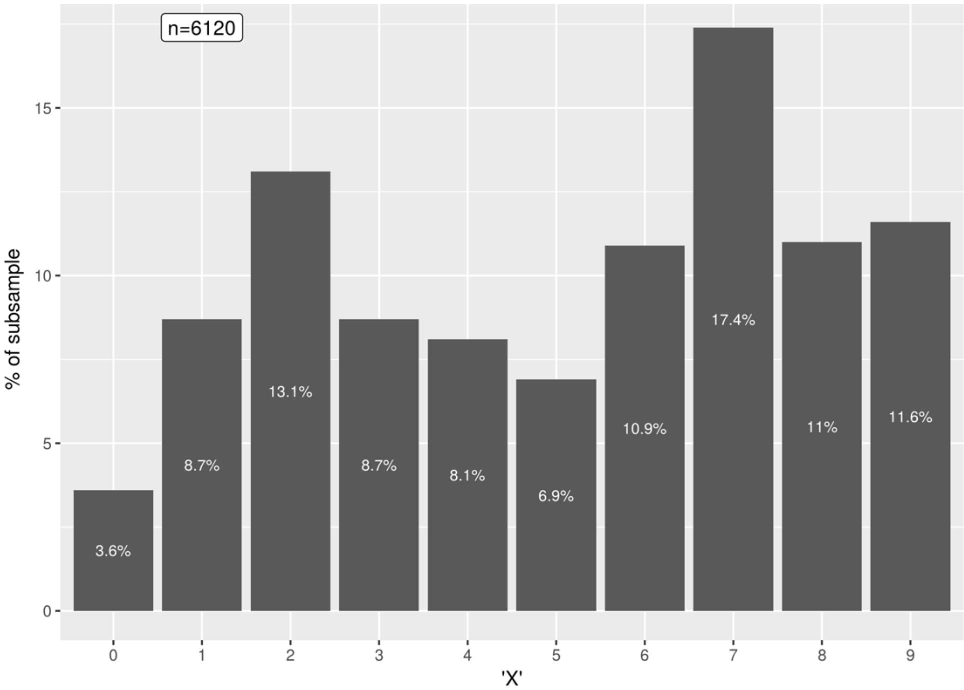

Figure 29 Distribution of price ending for housing sales subsample, model: dX5,000

4.5.3 Alternative explanations: price level

Given the methodology, one possible explanation for the observed effect is that it could be that the effect is driven by the observations with prices to the larger end of the subsample, as the effect of being precise in the fifth digit decreases relatively for more expensive real estate. For instance, a unit change in the fifth digit is on average 6% of the price for an estate with a price at 100,000 levels, where it is about 1% of the price for an estate with a price at 900,000 levels.

To check if this explanation is true, first, the correlation between the fifth and sixth digit is gauged in order to judge whether the fifth digit can predict the sixth. Figure 30 presents the association plot of contingency table of 5th and 6th digits in this subsample. With a more formal approach, treating the digits again as nominal variables, the Cramer's V measure between the fifth and sixth digit is 0.096, which can be interpreted as a small association (Cohen, 1988) as Cramer's V ranges from 0 to 1.

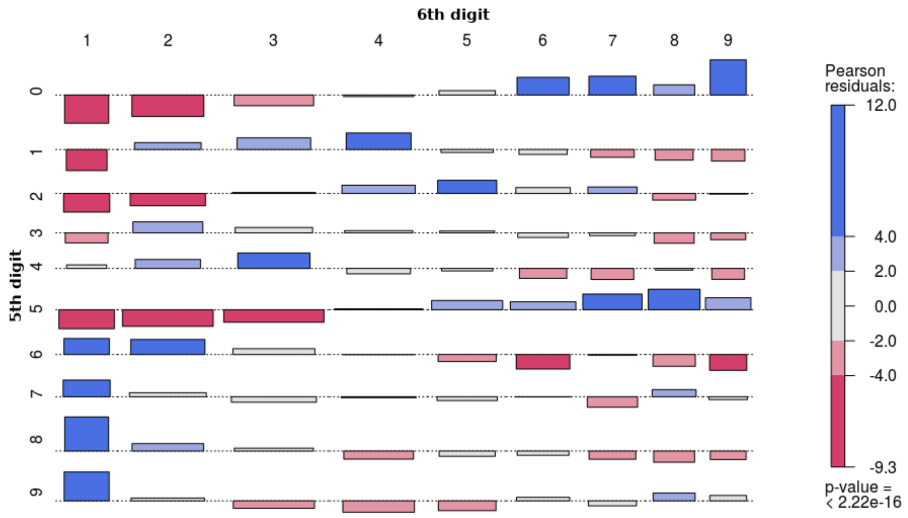

Figure 30 Association plot of contingency table of 5th and 6th digits in housing subsample with 6-digit price tags

Second, the secondary regression analysis is replicated using the ratio of residuals to prices instead of residuals, normalizing the effect of price levels. The predictive power and the significance of the results of the fifth digit on the residual ratios are similar to those without normalization (Table 28), indicating that the level of prices is not the full explanation of the phenomena.

Table 28. Using Residual Ratios instead of Residuals

	Dependent variable:
	residualsRatios
	OLS
digit_group0	0.008***
	(0.001)
digit_group1	0.0003
	(0.001)
digit_group2	0.004***
	(0.001)
digit_group3	0.003***
	(0.001)
digit_group4	0.002**
	(0.001)
digit_group5	0.009***
	(0.0005)
digit_group6	0.004***
	(0.001)
digit_group7	0.003***
	(0.001)
digit_group8	0.004***
	(0.001)
digit_group9	0.004***
	(0.001)
Constant	−0.003***
	(0.0002)
Observations	15,673
R^2	0.029
Adjusted R^2	0.028
Residual Std. Error	0.019 (df = 15662)
F Statistic	46.737*** (df = 10; 15662)
Note:	*p<0.1; **p<0.05; ***p<0.01

For further checking if the level of prices is in effect, another pricing model is formed, where two categorical variables are added to the initial regression: the 6th digit and the 5th digit group based on price-ending patterns. With doing so, it is possible to check the effect of the digit groups still has an effect after controlling for the first digit, when the price ending is even. The results, shown in Table 29 confirm that the clustering around even numbers is still statistically significant on pricing and the price level explanation is not sufficient to explain the observed effect of the even-pricing phenomena.

4.5.4 Considering truncation effects

Due to the methodology used, observations are restricted to include only 6 digit samples. This is a data slicing on the dependent variable and might cause biases in regressions; therefore the results are checked if the truncation have any effect on the inferences. Consider four truncated regressions models. The first one uses the pricing model given in Equation 60:

$$log(P)^{*} = X + u \tag{66}$$

The second truncated regression model adds the digit group variable D to the model:

$$log(P)^{*} = X + D + u \tag{67}$$

Table 29 Controlling for Price Levels, Housing Subsample with 6-digit Price-Tags, Independent Variables Omitted for Brevity

	Dependent variable:
	log(Price)
	OLS
...	...
	...
digit_group1	−0.110***
	(0.007)
digit_group2	−0.076***
	(0.006)
digit_group3	−0.056***
	(0.006)
digit_group4	−0.035***
	(0.007)
digit_group5	−0.006
	(0.004)
digit_group6	0.014**
	(0.006)
digit_group7	0.039***
	(0.007)
digit_group8	0.066***
	(0.006)
digit_group9	0.085***
	(0.006)
as.factor(price_last6th)2	0.426***
	(0.006)
as.factor(price_last6th)3	0.739***
	(0.007)
as.factor(price_last6th)4	0.978***
	(0.007)
as.factor(price_last6th)5	1.177***
	(0.008)
as.factor(price_last6th)6	1.337***
	(0.009)
as.factor(price_last6th)7	1.477***
	(0.009)
as.factor(price_last6th)8	1.598***
	(0.010)
as.factor(price_last6th)9	1.710***
	(0.011)
Constant	12.083***
	(0.032)
Observations	3,088
R^2	0.981
Adjusted R^2	0.981
Residual Std. Error	0.067 (df = 3003)
F Statistic	1,878.699*** (df = 84; 3003)

The fourth model adds the digit group variable to the third model:

$$log(P)^{*} = X + L + D + u \quad (69)$$

In both of the second and forth truncated regression, the digit group variable D is significant and in line with the corresponding non-truncated models. Formally, consider two 2-degrees-of-freedom chi-square tests between likelihoods of model 1 and 2, and between model 3 and 4. Both tests have a p-value of 0.00; indicating the significance of the added variable D. Further tables, including truncated regression results, are provided in Appendix E. As a conclusion, the truncation has no effect on the observed phenomena.

4.5.5 Extending analysis: observations with different digit counts

The methodology used in the analysis requires the price figures to have the same number of digits; however, not all observations have the same number of digits in their price figures. 75.6% of the prices in the housing market for sale dataset have 6-digit figures where 23.6% of the observations have 7-digit prices. Main analysis is focused on the 6-digit subsample as not only it constitutes the majority of the sample but also it covers the observations nicely with a median (490,000TRY) that is conveniently around the middle of 6 digit numbers, eliminating a possible cause of clustering purely due to the level of prices.

In this subsection, the analysis is replicated for the observations that have 7-digit price figures for housing markets for sale. The dataset is similarly subsampled to contain the observations that have 7 digit price figures. The last four digits of this subsample is predominantly ‘0000’ and there are some observations with an ending of ‘5000’ (Figure 31). Out of the observations with a price-ending of ‘0000’, the clustering at the 5th digit is more prevalent compared to the 6-digit subsample. More than two thirds of the observations are clustered around 0 and 5 in the 5th digit in total (Figure 32).

A pricing model is similarly formed, the group variable being again the fifth digit from the right, as in Figure 32. The model is applied to this sample of housing observations that have 7 digit price figures. Full regression specification and the results of the regression are given in Appendix D. The same analysis for the base case is applied to the 7 digit sample, looking at the effect of the 5th digit. The results, for which the tables are provided in Appendix E and the figures are provided in Appendix F, are overall similar to the results of the same analysis on the 6 digit sample. Particularly, the residual means in groups defined by the 5th digit are jointly statistically different from zero. The digit group variable is significant in predicting the residuals and the explanation powers are about the same (R^2 = 3.3%). The results are robust when residual ratios, calculated by dividing residuals by the log of prices, are used as the response variable, in an attempt to control for the effect of price levels.

Stepping a digit up, when we examine the distribution of the last 5 digits in this sample, the increments of ‘5000’ disappears and becomes the increments of ‘50000’ (Figure 33) similar to what is observed with the distribution of the last 4 digits in the 6 digit sample. The increments are sparser and more clustered, suggesting that the usage of the course pricing set is not due to being able to be precise about the house prices. If people can be precise about the value of a house

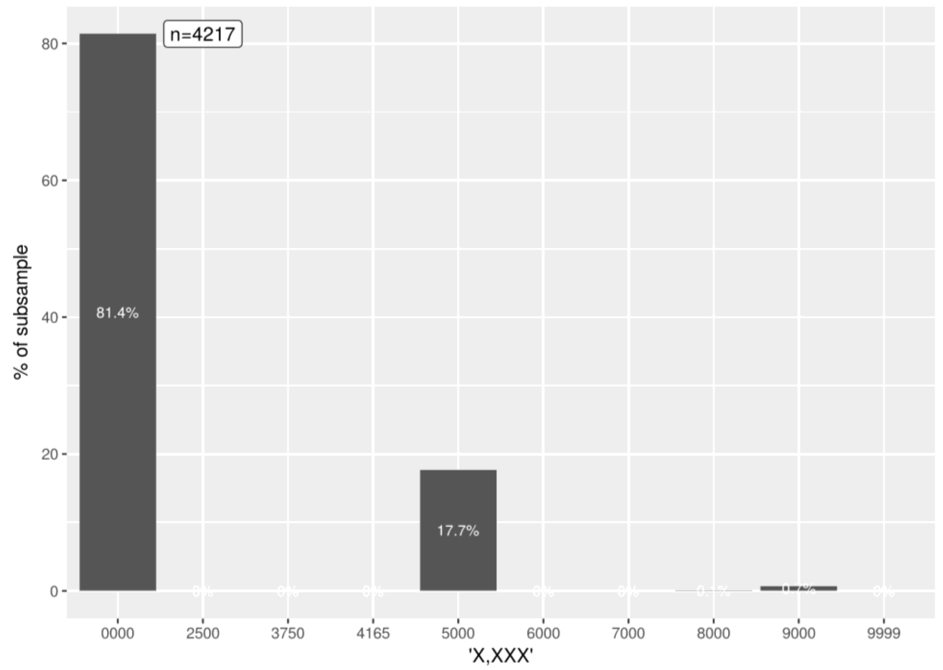

Figure 32 Distribution of price ending for housing sales subsample, model: d,ddX,XXX

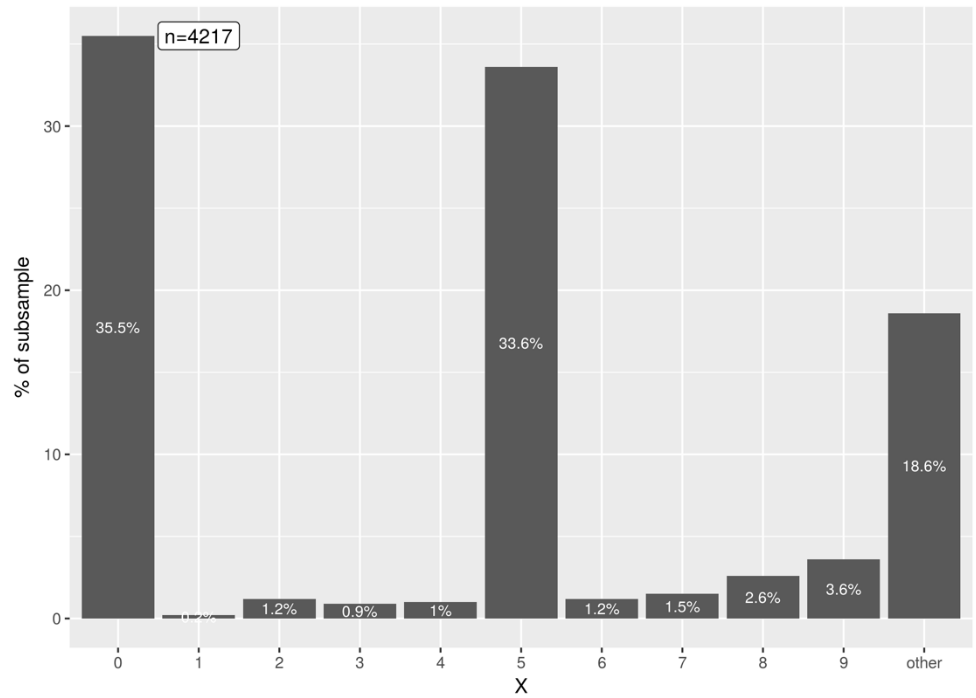

Figure 31 Distribution of price ending for housing sales subsample, model: d,dX0,000

with ‘5000’ precision in the 6-digit sample, then they could be as precise in the 7-digit sample as well. The effect is not likely to be due to increased marginal time cost of being precise due increased purchasing power either, as the stakes are so high that outsourcing the process of finding out a more precise price is profitable at this point. For instance, while we observe clear clustering at 50,000 increments, the minimum real estate property valuation fees in the market in 2017 as defined by the Capital Markets Board of Turkey is 445TL (Sermaye Piyasasi Kurulu, 2016). Thus, the effect must be due to some other cause than an information friction, such as a tendency to rely on even numbers, easing the negotiation process or precision avoidance such as for instance to signal luxury.

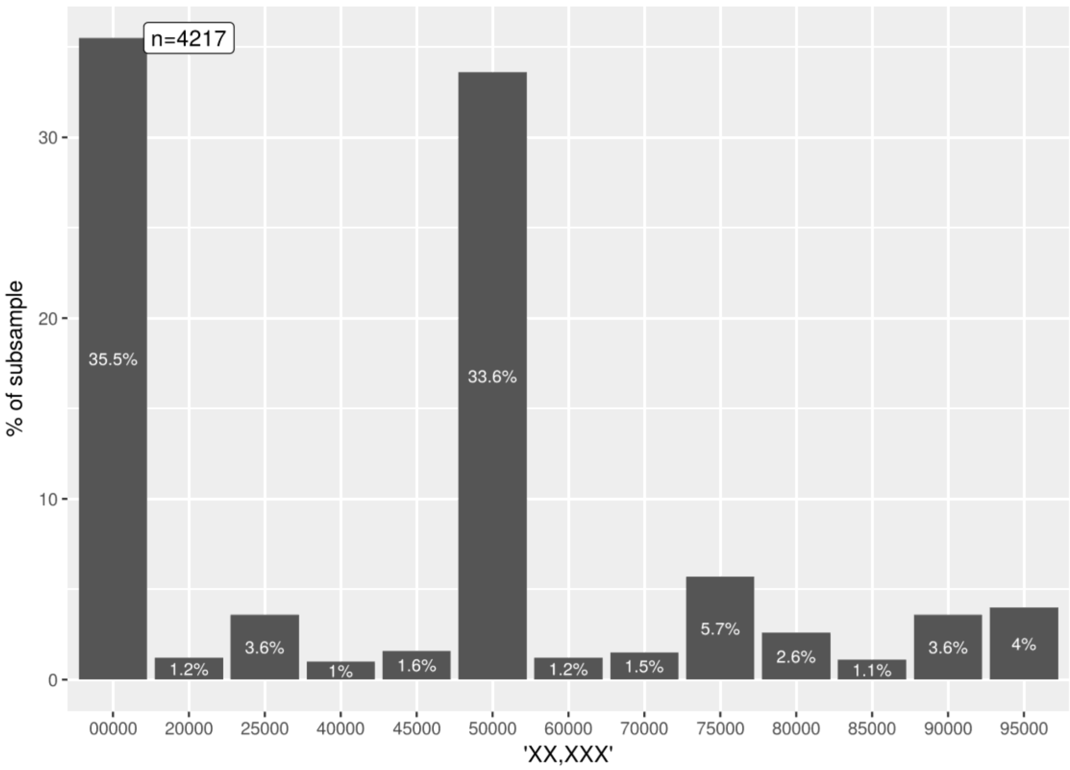

Figure 33 Distribution of price ending for housing sales subsample, model: d,dXX,XXX, bins < %1 excluded in the plot

The clustering around 0 and 5 in the second most important digit disappears in the housing market for sales with 7 digit prices, as shown in Figure 34. The ratio of the observations with remaining digits being all zeros, ‘00000’ in this case, decreases greatly with respect to ‘0000’ in a 6 digit analysis.

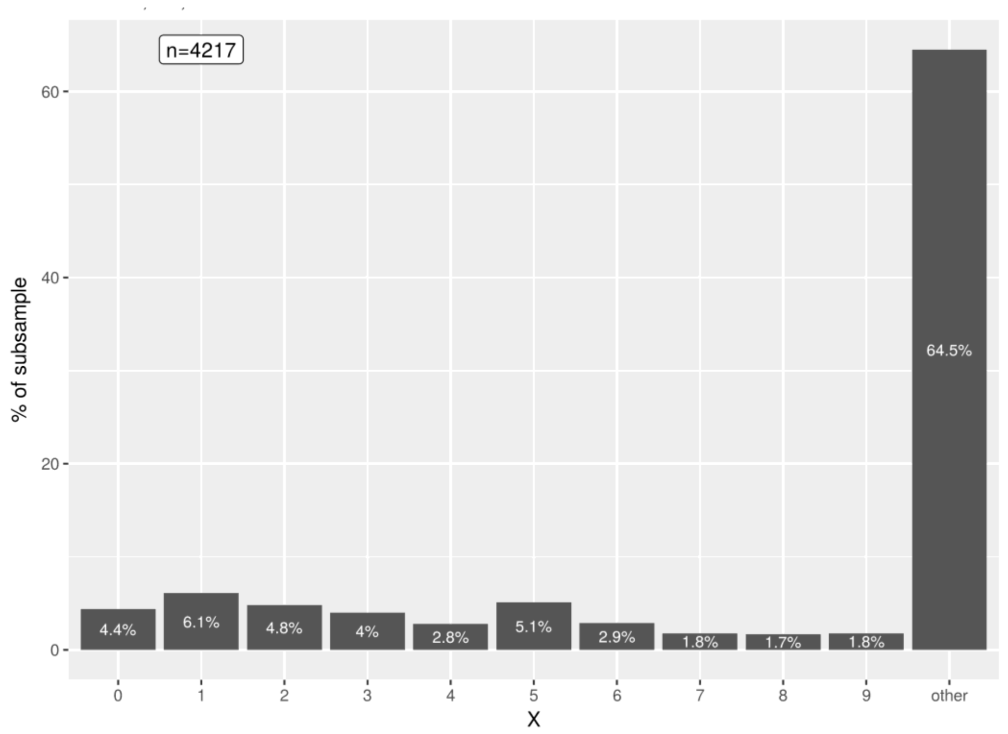

Figure 34 Distribution of price ending for housing sales subsample, model: d,X00,000

Regardless, the significance of the digit groups in explaining the residuals remain the same, albeit with different coefficients, the residual means across groups defined by the digit group variable are not jointly zero due to an ANOVA test with a 0.0000 significance, and the total explanation power of the digit count increases. The regression results are robust when residual ratios (standardized residuals by prices) are used and the results of the regression are presented in Appendix E.

4.5.6 Extending analysis: commercial real estate market

An important is whether the observed even pricing phenomena is particular to the housing market or it extends to other markets as well. In order to have empirical evidence, the study is extended to the commercial real estate market in this section and other markets in next sections. The methodology used for the housing market applied directly to the commercial real estate market.

Observations having prices with 6-digit figures constitute 56.7% of the data set for the commercial market for sales. 39.4% of the observations have 7 figured prices, which indicate somewhat higher price levels than the housing market. Figure 35 shows the clustering in the 6-digit subsample of commercial properties for sales, where clustering around even price-endings ‘00,000’ and ‘50,000’ is still observed and in fact it is more prevalent than in the housing market. The increments of ‘5000’

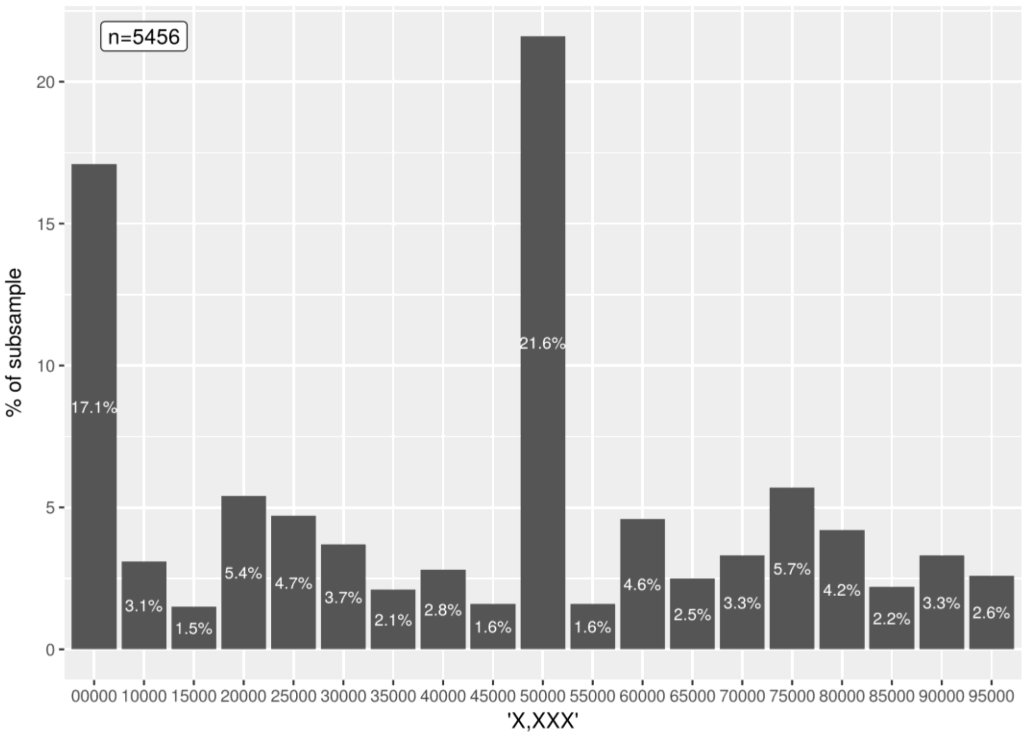

Figure 35 Distribution of price ending for commercial property sales subsample, model: dXX,XXX, bins < %1 excluded in the plot

in pricing is also observed similar to the housing market for sales, albeit with a little less frequency.

Even though the trace of increments of '5000' can be found in the data set, most observations in the subsample ends with '0,000' (Figure 36). Out of the 6-digit observations that end with '0,000', a clustering around 0 and 5 is observed (Figure 37) and major clustering happens around '00,000' and '50,000' in commercial real estate observations with 6 digit price tags. The phenomena observed is similar to those in the housing market, where the differences are that (a) there are fewer properties in the irregularly priced "other" bin and the even pricing tendency is higher. The differences in the markets suggest more precision avoidance in the commercial real estate market and favors the negotiation ease explanation as commercial real estate properties tend to be owned by owners with more experience in trade, compared to the housing market.

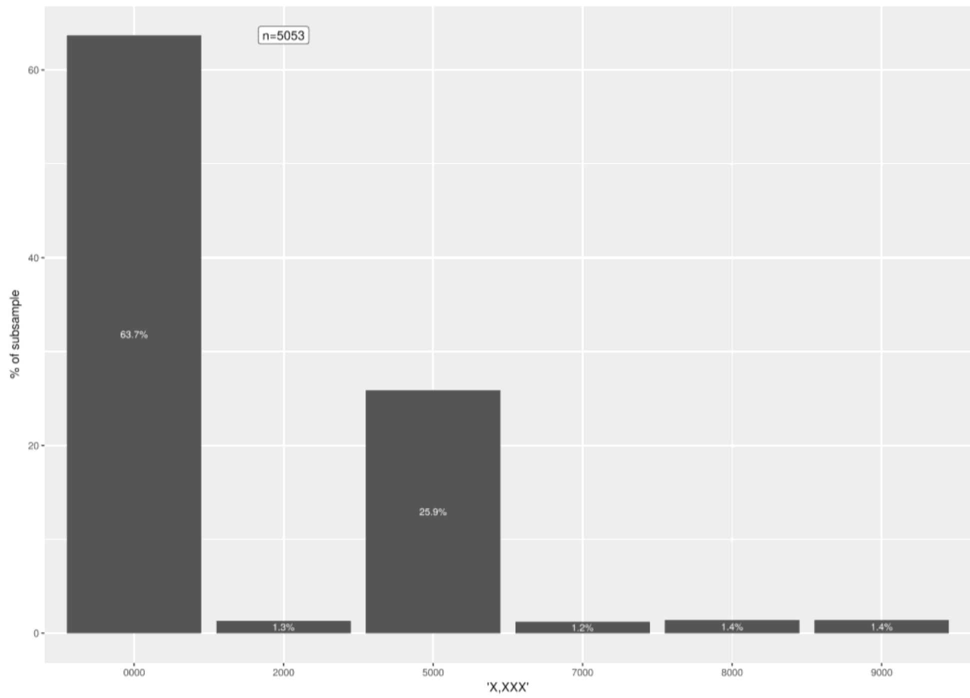

Figure 36 Distribution of price ending for commercial property sales subsample, model ddX,XXX, bins < %1 excluded in the plot

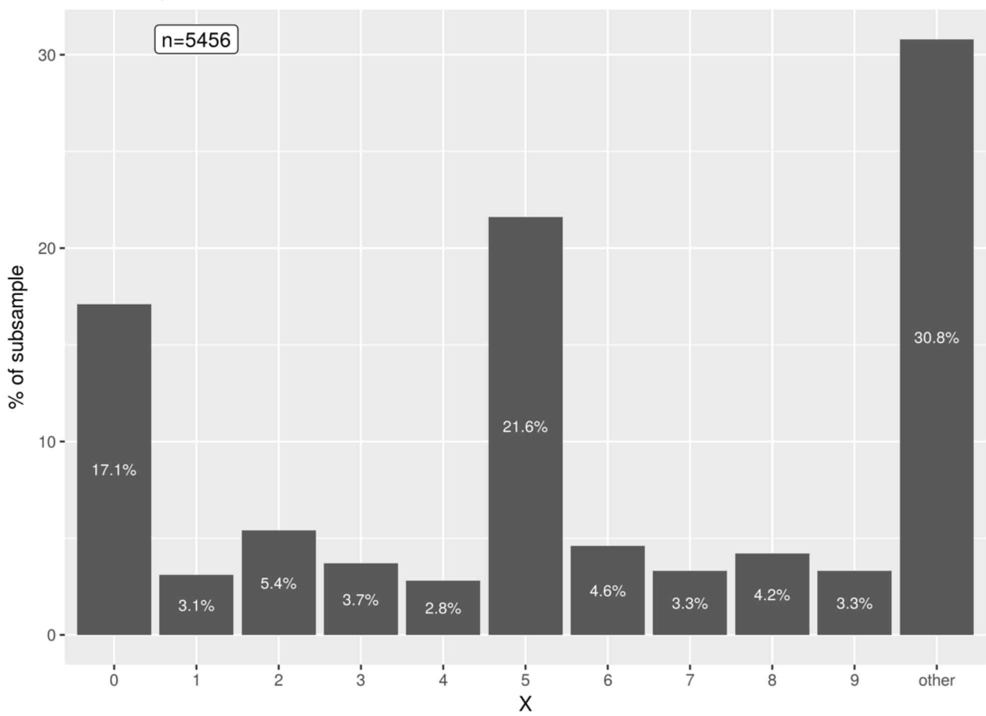

Figure 37 Distribution of price ending for commercial property sales subsample: model dX0,000

In commercial properties for sales with 7-digit price figures, the clustering around 0 and 5 is also in effect (Figure 38) for the second most important digit. This suggests that the clustering is not about the inability to be precise about the valuation, as the pricing patterns in 6 digit prices show up to be the same in 7 digit prices with a digit shifted. (Figure 36 and Figure 39). In contrast, this implies that the pricing patterns observed are more related to either cognition or communication of market agents.

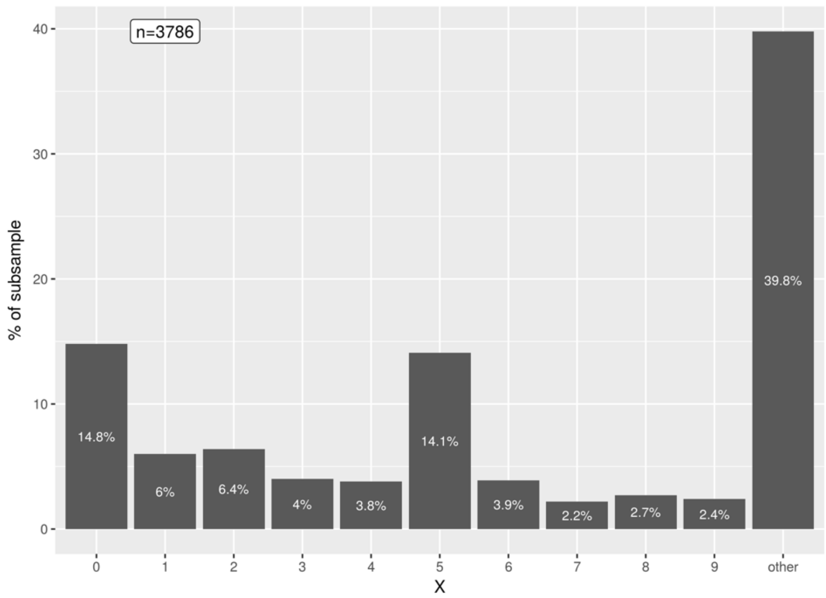

Figure 38 Distribution of price endings for commercial property sales subsample, model: d,X00,000

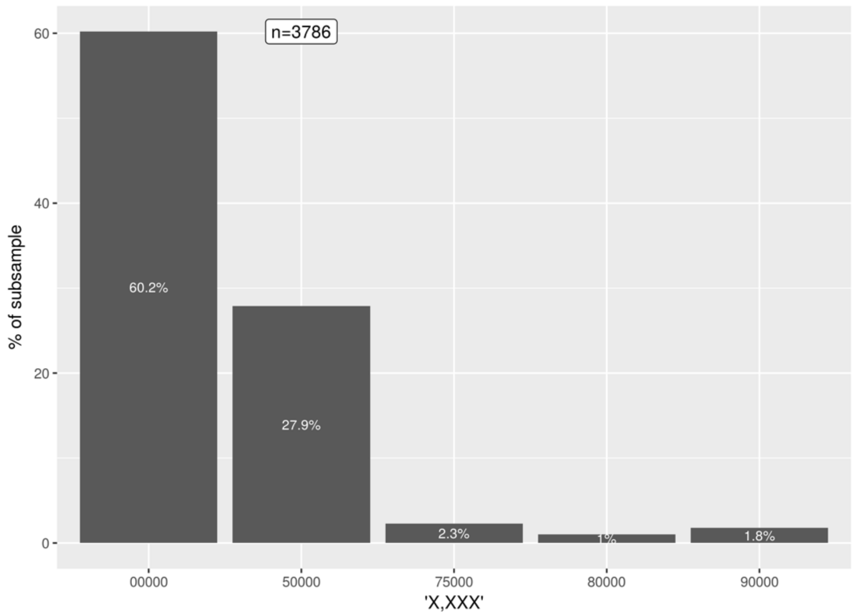

Figure 39 Distribution of price endings for commercial property sales subsample, model d,dXX,XXX, bins < %1 excluded in the plot

4.5.7 Extending analysis: land markets

As in line with housing and commercial markets for sales, land markets for sales convey a similar clustering in prices at the second most important digit (Figure 40). Particularly, the prices move in increments of '5,000' and there is clustering around price endings '00000' and '50000', at the expense of the increments next to them (Figure 41). As land, commercial and housing markets have different dynamics of pricing as well as different kinds of friction such as liquidity or different tax rates, similar findings suggest that even pricing is a behavioral phenomena regarding prices.

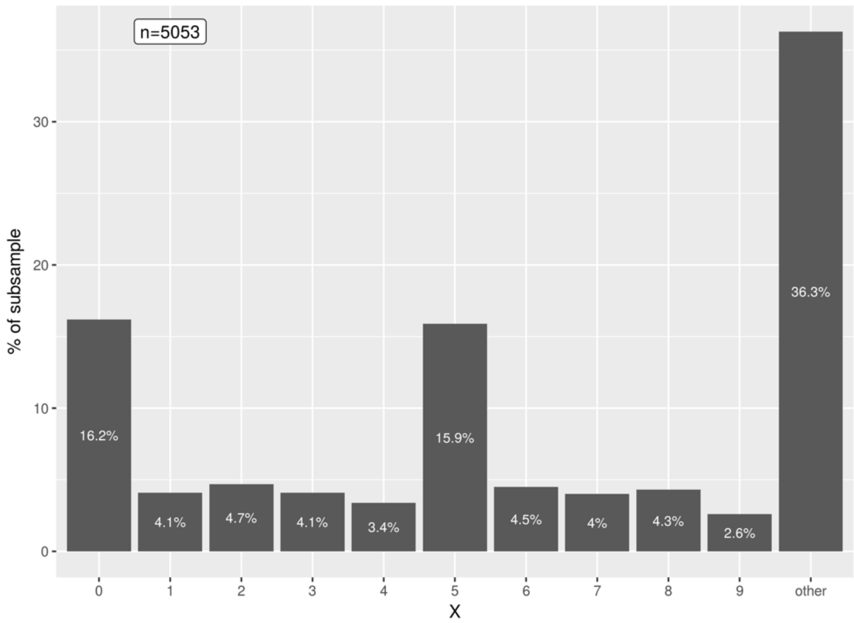

Figure 40 Distribution of price endings for land sales subsample, model: dX0,000

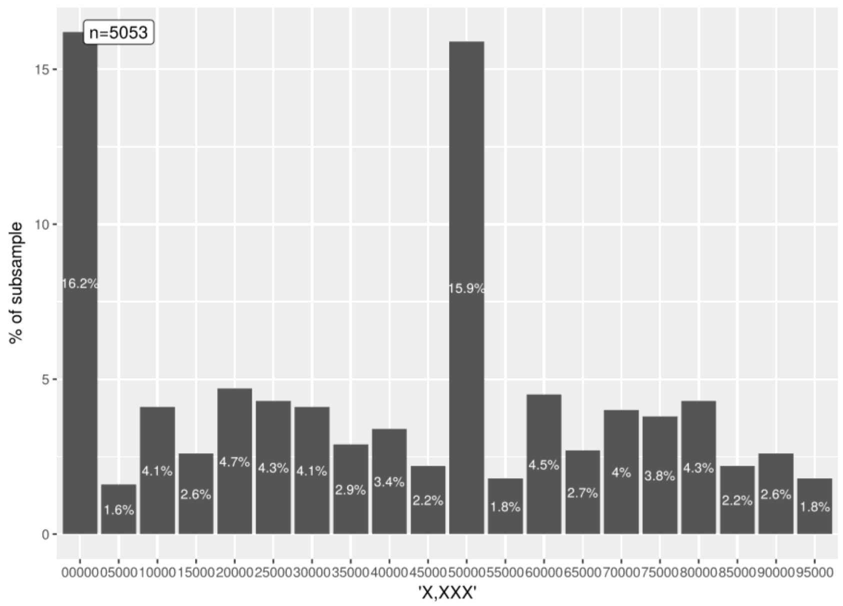

Figure 41 Distribution of price endings for land sales subsample, model dXX,XXX, bins < %1 excluded in the plot

4.5.8 Extending analysis: housing markets for rental

One rationale of this study is that the pricing patterns emerged as a result of pricing behavior can be economically significant when the transactions are high-level transactions as in real estate sales transactions. This premise is not likely to be valid for housing rentals as most of them do not realize at high stakes; regardless, the analysis in this domain gives us insights about what the observed phenomena is due to.

Since the price levels are different in rental markets, the methodology is adopted to 4 digits, the digit under scrutiny being again the second most important digit, that is the third digit instead of the fifth. The median rental price for all rental houses in the data set is 2200TRY. The analysis focuses on rentals with 4 digit price levels, which cover 93% of all observations.

When we look at the price ending distributions of all housing rentals with 4 digit prices, a similar pattern to previous real estate markets is seen. The price increments have the step size of '50' and there is a large clustering around '000' and '500' (Figure 42 and Figure 43). Out of the observations with a price ending of '00', if we focus on the second most important digit in a similar manner to the analysis done in previous sections, we can group the observations into bins. The relative bin sizes will be as shown in Figure 44.

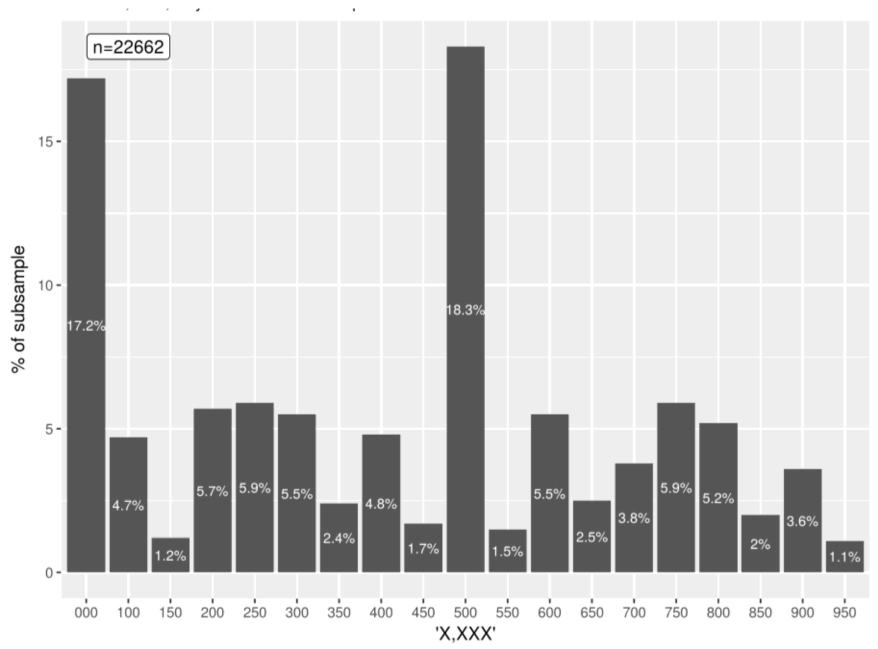

Figure 42 Distribution of price endings for housing rentals subsample, model d,XXX, bins < %1 excluded in the plot

The pricing model defined for housing sales, as given in Equation 60, is applied to the rental sample. The regression in the rental market results in a fit with an R^2 of 0.73, that is similar to the R^2 of the regression for housing sales. In order to check whether the clustering around even price figures have an effect, the residuals of the regression are regressed similarly to the digit groups shown in Figure 44.

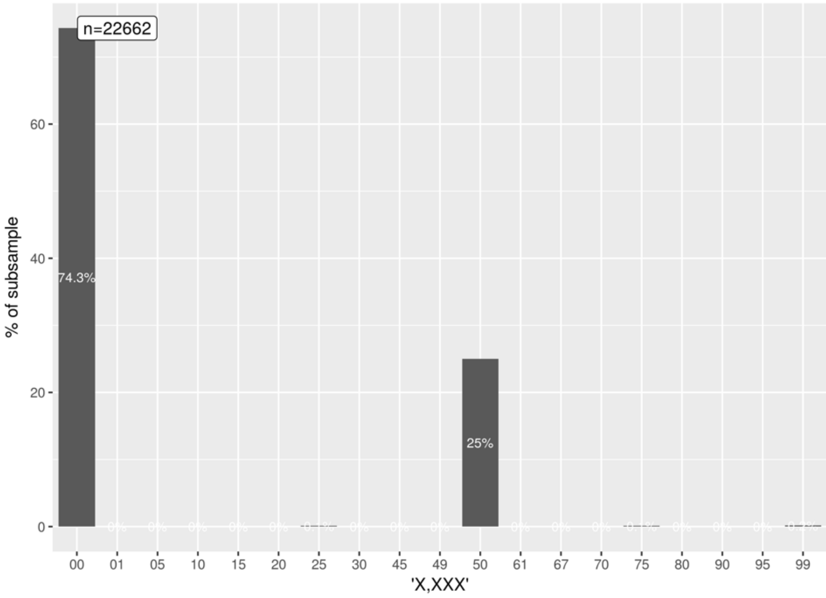

Figure 43 Distribution of price endings for housing rentals subsample, model d,dXX

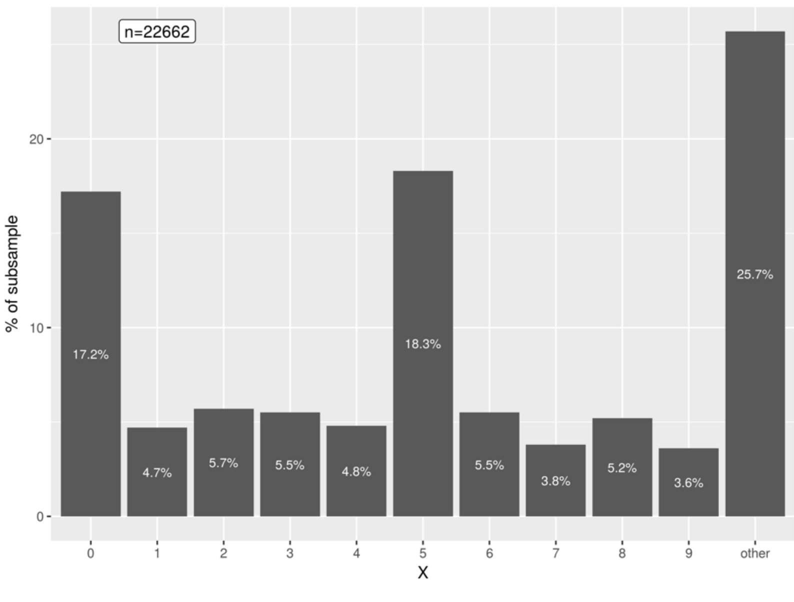

Figure 44 Distribution of price endings for housing rentals subsample, model d,X00

According to this regression, the digit groups can significantly explain about 2.9% of the variation in residuals of rental pricing regression. The results are presented in Appendix E. The results of regression explaining variation in pricing regression residuals are robust to effects of price levels, as checked by replacing the residual term (the response term) with residual divided by the log of rental price.

4.5.9 Comparison of results in the literature

Data of this study conveys comparable price-clustering patterns to previous reports in the literature. Palmon et al. (2004) study housing sales in a suburban district in Texas (USA) and Beracha and Seiler (2014) study housing sales in urban areas in Virginia (USA); while this paper study housing sales, commercial sales, land sales, and housing rental data across Istanbul (Turkey). A comparison of price clustering patterns is these studies are given in Table 30.

Table 30. Comparison of Even Pricing Tendency in Different Markets

	Current Study Housing Sales	Commercial	Land	Housing Rental	Beracha and Seiler (2014)	Palmon et al. (2004)
0	13.7%	24.7%	25.5%	23.1%	14.3%	19.6%
1	7.5%	4.5%	6.4%	6.4%	2.7%	
2	9.2%	7.9%	7.4%	7.6%	5.0%	
3	8.8%	5.4%	6.4%	7.3%	3.5%	
4	6.5%	4.0%	5.4%	6.5%	9.7%	
5	21.2%	31.2%	24.9%	24.7%	16.3%	17.2%
6	9.2%	6.7%	7.0%	7.4%	3.5%	
7	7.3%	4.8%	6.2%	5.1%	4.9%	
8	8.7%	6.1%	6.7%	7.0%	4.9%	
9	7.8%	4.8%	4.0%	4.9%	35.4%	51.9%

The comparison yields similarity in even pricing around 0 and 5 in multiple real estate markets, across cities and across cultures. Importantly, however, charm pricing is not observed in Istanbul, while both of the other studies report very significant charm pricing behavior in real estate. The difference suggests that there is a dominant cultural aspect of price clustering, validating the plausibility of pricing phenomena to be behavioral.

4.6 Discussion and conclusion

In this study, it is documented that there is a price clustering pattern in various real estate markets that it is formally shown that it implies an anomaly in pricing. The price clustering pattern results in particular price endings being more prevalent than others in real estate listings. Particularly, the use of even endings such as with zeros and the increments of fives are more evident in the data than a normally or lepkokurtically distributed price data or uniformly distributed price ending data would have predicted. This pricing behavior is comparable to the previous studies such as Palmon et al. (2004) and Beracha and Seiler's (2014) which both report clustering on zeros and fives.

The even pricing around 0 and 5 occurs at multiple levels of digits, meaning that while a price tag of '300.000' is more common than '310,000', the price tag of '310,000' is also more common than '311,000$' and so on. Clustering around 5 is more prevalent than clustering around 0, and the increments of '5000' can be detected in data. The clustering happens at the expense of the digits next to them, usually 1, 4, and 6, and a "precision avoidance" is observed. For instance a price ending of '05000' (or '050' in rentals) is distinctively unpopular.

The economic implications of the price-ending behavior is examined through an analysis of variance on the hedonic pricing model. The findings show that the digits of the listing prices are statistically significant in the pricing model, although, any digit sequence in prices should not have any effect on the price of a property. Moreover, the average rounding margin is not only comparable to the median disposable income but also higher than the expert valuation costs. That is, the rounding margin, conservatively accepted as one tenth of the second most important digit or one percent of the price, for the median house in the dataset of this study is 5,300TL while the monthly disposable median income in Turkey is 1,324TL. It is also observed that the rounding margin is higher than the average expert valuation of a housing property, which is against the information inefficiency explanation for the rounding behavior.

We find that the even-pricing behavior exists in multiple real estate markets. In addition to the housing market, even-pricing exists in commercial and land markets where liquidity and market participant information levels might be different. Next, it is shown that even-pricing exists in the rental markets as well, where the pricing levels and respective economic impact vary. Besides, findings of this study confirm that even-pricing phenomena exist in the Turkish market but with a cultural difference on charm pricing. Specifically, while charm pricing is observed in the US housing market in multiple states, it is not observed in the Istanbul real estate markets.

Findings of this study support the cognition and negotiation hypotheses in the literature on price clustering while alternative frictional explanations such as taxes or information inefficiency remain insufficient in explaining the evidence. The observed effect continues after controlling for price levels in multiple real estate markets.

Information inefficiency cannot be the reason because (a) same even pricing behavior is observed in 6-digit and 7-digit levels and (b) in rentals, (c) real estate transactions are high stake transactions and (d) it is clearly profitable to outsource valuation at 7-digit levels. Tax frictions cannot be the reason because, same even pricing behavior occurs in housing markets, commercial markets, and land markets as well as rental markets where pricing dynamics and taxes are different. Moreover similar behavior is observed across countries and cities, which differ in their tax considerations. Findings of this study on the other hand is consistent with cognitive explanations as increments using the figures 0 and 5 is observed consistently across levels, markets and cultures. We also detect "precision avoidance" behavior, where the actors seem to shy away from price tags that seem too precise. (a) The precision avoidance behavior, (b) even pricing occurring at the expense of neighboring digits and (c) particular differences between commercial and housing markets are in line negotiation process explanations. Finally, it is documented that there is a cultural aspect in real estate pricing, since charm pricing is not observed in the data set of this study.

APPENDIX A

FULL LIST OF MODELS AND VARIABLES

		Model 1	Model 2	Model 3	Model 4
Dependent					
log(price)		✓	✓	✓	✓
Independent					
(Building Related)					
property_type	Type of housing such as flat, residence or yali dairesi. Categorical.	✓	✓	✓	✓
age	Age of the building. Numeric.	✓	✓	✓	✓
heating_type	Type of heating. Categorical.	✓	✓	✓	✓
construction_type	Type of construction. Categorical.	✓	✓	✓	✓
gated	Whether the building belongs to a gated community. Binary.	✓	✓	✓	✓
(Apartment Related)					
area	Area of the unit. Numeric.	✓	✓	✓	✓
bathrooms	Number of bathrooms in the unit. Numeric.	✓	✓	✓	✓
living_rooms	Number of living rooms in the unit. Numeric.	✓	✓	✓	✓
bedrooms	Number of bedrooms in the unit. Numeric.	✓	✓	✓	✓
floor_at_clustered	Height of the unit in the building in floors. Categorical.	✓	✓	✓	✓
(Neighborhood Related)					
district	Dummies that correspond to 39 districts of Istanbul.	✓	✓		
district_neighborhood	Dummies that correspond to 517 neighborhoods of Istanbul.			✓	✓

(Geo-Locational)			
dist.2.closest.fire_department	Distance of the unit to the closest fire department in meters. Numeric.	✓	✓
dist.2.closest.hospital	Distance of the unit to the closest hospital in meters. Numeric.	✓	✓
dist.2.closest.shopping_mall	Distance of the unit to the closest shopping mall in meters. Numeric.	✓	✓
dist.2.closest.police_department	Distance of the unit to the closest police department in meters. Numeric.	✓	✓
dist.2.closest.transportation	Distance of the unit to the closest subway or metrobus stop. Numeric.	✓	✓
dist.2.closest.private.preschool	Distance of the unit to the closest private preschool. Numeric.	✓	✓
dist.2.closest.private.primschool	Distance of the unit to the closest private primary school. Numeric.	✓	✓
dist.2.closest.private.highschool	Distance of the unit to the closest private high school. Numeric.	✓	✓
dist.2.closest.private.middleschool	Distance of the unit to the closest private middle school in meters. Numeric.	✓	✓
dist.2.closest.public.preschool	Distance of the unit to the closest public preschool in meters. Numeric.	✓	✓
dist.2.closest.public.primschool	Distance of the unit to the closest public primary school in meters. Numeric.	✓	✓
dist.2.closest.public.highschool	Distance of the unit to the closest public high school in meters. Numeric.	✓	✓
dist.2.closest.public.middleschool	Distance of the unit to the closest public middle school in meters. Numeric.	✓	✓
dist.2.closest.university	Distance of the unit to the closest university campus in meters. Numeric.	✓	✓

APPENDIX B

FULL REGRESSION RESULTS – CHAPTER 3, ALL MODELS

	Dependent variable:			
	log(price)			
	(1)	(2)	(3)	(4)
Constant	13.117***	13.155***	13.796***	13.209***
	(0.111)	(0.117)	(0.219)	(0.233)
age	−0.003***	−0.003***	−0.007***	−0.007***
	(0.0003)	(0.0002)	(0.0002)	(0.0002)
area	0.008***	0.008***	0.007***	0.007***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
bathrooms	0.123***	0.109***	0.079***	0.077***
	(0.005)	(0.005)	(0.004)	(0.004)
living_rooms	−0.186***	−0.165***	−0.108***	−0.107***
	(0.010)	(0.010)	(0.008)	(0.008)
bedrooms	−0.026***	−0.012***	0.013***	0.015***
	(0.004)	(0.004)	(0.004)	(0.004)
heating_typeKat Kaloriferi	−0.234***	−0.121*	−0.080	−0.080
	(0.066)	(0.065)	(0.069)	(0.068)
heating_typeKlima	−0.517***	−0.392***	−0.364***	−0.341***
	(0.065)	(0.064)	(0.068)	(0.068)
heating_typeKombi	−0.458***	−0.336***	−0.276***	−0.261***
	(0.052)	(0.052)	(0.060)	(0.060)
heating_typeMerkezi	−0.346***	−0.227***	−0.195***	−0.180***
	(0.052)	(0.052)	(0.061)	(0.060)
heating_typeMerkezi (Pay Ölçer)	−0.362***	−0.247***	−0.177***	−0.165***
	(0.052)	(0.052)	(0.060)	(0.060)
heating_typeSoba	−0.662***	−0.532***	−0.379***	−0.367***
	(0.065)	(0.064)	(0.069)	(0.068)
heating_typeYok	−0.667***	−0.522***	−0.393***	−0.381***
	(0.091)	(0.089)	(0.086)	(0.085)
construction_typeBetonarme	0.043	0.022	−0.029	−0.018
	(0.083)	(0.081)	(0.067)	(0.066)
construction_typeÇelik	0.122	0.126	0.080	0.091
	(0.089)	(0.087)	(0.072)	(0.071)
construction_typeKagir	0.271***	0.257***	0.108	0.113
	(0.095)	(0.092)	(0.076)	(0.075)
construction_typeKütük	−0.106	−0.105	0.086	0.119
	(0.263)	(0.256)	(0.211)	(0.209)
construction_typePrefabrik	0.714**	0.812**	0.501*	0.595**
	(0.363)	(0.353)	(0.287)	(0.285)
construction_typeTaş Bina	0.598**	0.403	0.419**	0.451**
	(0.263)	(0.256)	(0.209)	(0.207)
construction_typeYığma	0.178	0.149	−0.121	−0.103
	(0.133)	(0.129)	(0.111)	(0.110)
property_typeLoft Daire	−0.150	−0.120	−0.074	−0.065
	(0.113)	(0.110)	(0.094)	(0.093)
property_typeResidence	0.054***	0.080***	0.145***	0.145***
	(0.012)	(0.012)	(0.010)	(0.010)
property_typeYalı Dairesi	1.470***	1.490***	1.010***	0.989***
	(0.046)	(0.045)	(0.041)	(0.040)
storey_at_clusteredgroundLevel	−0.168***	−0.175***	−0.170***	−0.172***
	(0.009)	(0.009)	(0.007)	(0.007)
storey_at_clusteredhalfStorey	−0.147***	−0.137***	−0.111***	−0.111***
	(0.011)	(0.010)	(0.009)	(0.009)
storey_at_clusteredhighStorey	0.071***	0.075***	0.070***	0.068***
	(0.009)	(0.009)	(0.008)	(0.007)
storey_at_clusteredmiddleStorey	0.023***	0.024***	0.031***	0.030***
	(0.007)	(0.007)	(0.006)	(0.006)
storey_at_clusteredtopLevel	−0.143***	−0.134***	−0.089***	−0.088***
	(0.009)	(0.009)	(0.007)	(0.007)
storey_at_clusteredundergroundLevel	−0.383***	−0.373***	−0.316***	−0.320***
	(0.019)	(0.019)	(0.016)	(0.016)
storey_at_clusteredveryHighStorey	0.083***	0.098***	0.139***	0.141***
	(0.015)	(0.014)	(0.012)	(0.012)
gatedHayır	−0.161***	−0.143***	−0.120***	−0.117***
	(0.008)	(0.008)	(0.007)	(0.007)
dist.2.closest.fire_department		0.00001***		−0.00001**
		(0.00000)		(0.00001)
dist.2.closest.hospital		0.00000		0.00003***
		(0.00000)		(0.00001)
dist.2.closest.shopping_mall		0.00001*		0.00001*
		(0.00000)		(0.00001)
dist.2.closest.police_department		−0.00000		0.00002***
		(0.00000)		(0.00001)
dist.2.closest.transportation		0.00000		−0.00003***

		(0.00000)	(0.00001)
dist.2.closest.private.anaokulu		-0.00004^{***}	0.00002^{*}
		(0.00001)	(0.00001)
dist.2.closest.private.ilkokul		0.00002^{***}	0.0001^{***}
		(0.00001)	(0.00001)
dist.2.closest.private.lise		-0.00004^{***}	-0.0001^{***}
		(0.00001)	(0.00001)
dist.2.closest.private.ortaokul		-0.00005^{***}	-0.0001^{***}
		(0.00001)	(0.00001)
dist.2.closest.public.anaokulu		0.00000	-0.00003^{***}
		(0.00000)	(0.00001)
dist.2.closest.public.ilkokul		0.00002^{*}	-0.00000
		(0.00001)	(0.00001)
dist.2.closest.public.lise		-0.00001	0.00000
		(0.00001)	(0.00001)
dist.2.closest.public.ortaokul		-0.00003^{**}	0.00002^{*}
		(0.00001)	(0.00001)
dist.2.closest.university		0.0001^{***}	0.0001^{***}
		(0.00000)	(0.00001)
districtArnavutköy	-0.828^{***}	-1.460^{***}	
	(0.084)	(0.099)	
districtAtaşehir	-0.226^{***}	-0.403^{***}	
	(0.052)	(0.065)	
districtAvcılar	-0.864^{***}	-1.058^{***}	
	(0.052)	(0.066)	
districtBağcılar	-0.608^{***}	-0.872^{***}	
	(0.053)	(0.067)	
districtBahçelievler	-0.451^{***}	-0.663^{***}	
	(0.050)	(0.064)	
districtBakırköy	0.422^{***}	0.234^{***}	
	(0.051)	(0.065)	
districtBaşakşehir	-0.709^{***}	-1.007^{***}	
	(0.052)	(0.065)	
districtBayrampaşa	-0.434^{***}	-0.587^{***}	
	(0.060)	(0.071)	
districtBeşiktaş	0.670^{***}	0.538^{***}	
	(0.050)	(0.064)	
districtBeykoz	0.212^{***}	-0.055	
	(0.070)	(0.081)	
districtBeylikdüzü	-0.910^{***}	-1.150^{***}	
	(0.050)	(0.064)	
districtBeyoğlu	0.138^{***}	0.005	
	(0.053)	(0.066)	
districtBüyükçekmece	-0.682^{***}	-1.085^{***}	
	(0.053)	(0.067)	
districtÇatalca	-0.961^{***}	-0.820^{**}	
	(0.356)	(0.361)	
districtÇekmeköy	-0.550^{***}	-0.907^{***}	
	(0.055)	(0.069)	
districtEsenler	-0.658^{***}	-0.893^{***}	
	(0.061)	(0.073)	
districtEsenyurt	-1.063^{***}	-1.300^{***}	
	(0.051)	(0.065)	
districtEyüp	-0.317^{***}	-0.680^{***}	
	(0.050)	(0.067)	
districtFatih	-0.229^{***}	-0.375^{***}	
	(0.052)	(0.066)	
districtGaziosmanpaşa	-0.471^{***}	-0.646^{***}	
	(0.062)	(0.074)	
districtGüngören	-0.436^{***}	-0.651^{***}	
	(0.063)	(0.074)	
districtKadıköy	0.221^{***}	0.009	
	(0.050)	(0.064)	
districtKağıthane	-0.387^{***}	-0.556^{***}	
	(0.051)	(0.064)	
districtKartal	-0.461^{***}	-0.686^{***}	
	(0.051)	(0.065)	
districtKüçükçekmece	-0.553^{***}	-0.750^{***}	
	(0.050)	(0.064)	
districtMaltepe	-0.278^{***}	-0.470^{***}	
	(0.050)	(0.065)	
districtPendik	-0.671^{***}	-0.905^{***}	
	(0.053)	(0.067)	
districtSancaktepe	-0.870^{***}	-1.144^{***}	
	(0.055)	(0.066)	
districtSarıyer	0.291^{***}	0.149^{**}	

	(0.051)	(0.065)		
districtŞile	−0.346***	1.180***		
	(0.117)	(0.194)		
districtSilivri	−0.960***	−1.124***		
	(0.065)	(0.103)		
districtŞişli	0.111**	−0.028		
	(0.050)	(0.064)		
districtSultanbeyli	−0.792***	−1.187***		
	(0.066)	(0.076)		
districtSultangazi	−0.763***	−1.007***		
	(0.104)	(0.109)		
districtTuzla	−0.569***	−0.775***		
	(0.054)	(0.071)		
districtÜmraniye	−0.459***	−0.685***		
	(0.052)	(0.068)		
districtÜsküdar	0.008	−0.145**		
	(0.050)	(0.064)		
districtZeytinburnu	−0.223***	−0.385***		
	(0.056)	(0.069)		
neighAdalar_Büyükada-Maden			−0.690***	0.142
			(0.206)	(0.213)
neighAdalar_Büyükada-Nizam			−0.648***	0.129
			(0.207)	(0.213)
neighArnavutköy_Anadolu			−1.660***	−1.815***
			(0.235)	(0.247)
neighArnavutköy_Arnavutköy Merkez			−1.816***	−1.858***
			(0.281)	(0.290)
neighArnavutköy_Bolluca			−1.376***	−1.799***
			(0.230)	(0.243)
neighArnavutköy_Hadımköy			−1.673***	−1.452***
			(0.235)	(0.252)
neighArnavutköy_Hastane			−1.686***	−1.451***
			(0.281)	(0.294)
neighArnavutköy_Hicret			−1.545***	−1.868***
			(0.343)	(0.349)
neighArnavutköy_Mavigöl			−1.550***	−2.003***
			(0.256)	(0.268)
neighArnavutköy_Mustafa Kemal Paşa			−1.777***	−1.833***
			(0.343)	(0.351)
neighArnavutköy_Taşoluk			−1.836***	−1.894***
			(0.281)	(0.294)
neighAtaşehir_Aşık Veysel			−1.280***	−0.765***
			(0.243)	(0.254)
neighAtaşehir_Ataşehir Atatürk			−0.786***	−0.260
			(0.201)	(0.216)
neighAtaşehir_Barbaros			−0.719***	−0.230
			(0.202)	(0.216)
neighAtaşehir_Esatpasa			−1.223***	−0.652***
			(0.207)	(0.220)
neighAtaşehir_Fetih			−0.962***	−0.437*
			(0.230)	(0.241)
neighAtaşehir_İçerenköy			−1.151***	−0.636***
			(0.203)	(0.218)
neighAtaşehir_İnönü			−1.195***	−0.732***
			(0.208)	(0.221)
neighAtaşehir_Kayışdağı			−1.263***	−0.765***
			(0.204)	(0.217)
neighAtaşehir_Küçükbakkalköy			−0.916***	−0.411*
			(0.205)	(0.220)
neighAtaşehir_Mevlana			−0.968***	−0.557**
			(0.243)	(0.253)
neighAtaşehir_Mustafa Kemal			−1.163***	−0.559
			(0.343)	(0.349)
neighAtaşehir_Örnek			−1.111***	−0.510**
			(0.203)	(0.216)
neighAtaşehir_Yeni Çamlıca			−1.127***	−0.755**
			(0.343)	(0.348)
neighAtaşehir_Yenişehir			−0.893***	−0.344
			(0.205)	(0.220)
neighAvcılar_Ambarlı			−1.504***	−1.003***
			(0.210)	(0.224)
neighAvcılar_Cihangir			−1.640***	−1.120***
			(0.202)	(0.216)
neighAvcılar_Denizköşkler			−1.406***	−0.990***
			(0.203)	(0.219)
neighAvcılar_Firuzköy			−2.057***	−1.464***

	(0.244)	(0.255)
neighAvcılar_Gümüşpala	-1.582^{***}	-1.173^{***}
	(0.210)	(0.225)
neighAvcılar_Merkez	-1.640^{***}	-1.133^{***}
	(0.204)	(0.218)
neighAvcılar_Mustafa Kemal Paşa	-1.734^{***}	-1.122^{***}
	(0.201)	(0.215)
neighAvcılar_Tahtakale	-1.483^{***}	-1.160^{***}
	(0.205)	(0.218)
neighAvcılar_Üniversite	-1.401^{***}	-0.891^{***}
	(0.244)	(0.254)
neighAvcılar_Yeşilkent	-1.472^{***}	-1.021^{***}
	(0.281)	(0.287)
neighBağcılar_100. Yıl	-1.844^{***}	-1.320^{***}
	(0.226)	(0.237)
neighBağcılar_Bağlar	-1.400^{***}	-1.057^{***}
	(0.214)	(0.227)
neighBağcılar_Barbaros	-1.532^{***}	-1.186^{***}
	(0.211)	(0.225)
neighBağcılar_Çınar	-1.554^{***}	-1.114^{***}
	(0.281)	(0.290)
neighBağcılar_Demirkapı	-1.444^{***}	-0.893^{***}
	(0.220)	(0.234)
neighBağcılar_Evren	-1.333^{***}	-0.881^{***}
	(0.205)	(0.220)
neighBağcılar_Fatih	-1.407^{***}	-0.871^{***}
	(0.235)	(0.246)
neighBağcılar_Fevziçakmak	-1.173^{***}	-0.668^{***}
	(0.243)	(0.253)
neighBağcılar_Göztepe	-1.352^{***}	-0.753^{***}
	(0.205)	(0.220)
neighBağcılar_Güneşli	-1.299^{***}	-0.836^{***}
	(0.203)	(0.217)
neighBağcılar_Hürriyet	-1.361^{***}	-0.928^{***}
	(0.206)	(0.220)
neighBağcılar_İnönü	-1.267^{***}	-0.874^{***}
	(0.235)	(0.247)
neighBağcılar_Kazımkarabekir	-1.526^{***}	-1.149^{***}
	(0.211)	(0.225)
neighBağcılar_Kemalpaşa	-1.629^{***}	-1.092^{***}
	(0.343)	(0.349)
neighBağcılar_Kirazlı	-1.387^{***}	-0.938^{***}
	(0.212)	(0.226)
neighBağcılar_Mahmutbey	-1.147^{***}	-0.633^{***}
	(0.206)	(0.220)
neighBağcılar_Merkez	-1.301^{***}	-0.828^{***}
	(0.215)	(0.229)
neighBağcılar_Sancaktepe	-1.460^{***}	-1.014^{***}
	(0.281)	(0.290)
neighBağcılar_Yavuz Selim	-1.537^{***}	-1.047^{***}
	(0.218)	(0.230)
neighBağcılar_Yenigün	-1.391^{***}	-0.965^{***}
	(0.343)	(0.350)
neighBağcılar_Yenimahalle	-1.372^{***}	-0.859^{***}
	(0.220)	(0.234)
neighBağcılar_Yıldıztepe	-1.428^{***}	-1.063^{***}
	(0.212)	(0.225)
neighBahçelievler_Bahçelievler	-0.619^{***}	-0.093
	(0.201)	(0.215)
neighBahçelievler_Çobançeşme	-1.352^{***}	-0.865^{***}
	(0.204)	(0.218)
neighBahçelievler_Cumhuriyet	-1.354^{***}	-0.941^{***}
	(0.204)	(0.216)
neighBahçelievler_Fevziçakmak	-1.377^{***}	-0.859^{***}
	(0.213)	(0.226)
neighBahçelievler_Hürriyet	-1.405^{***}	-0.923^{***}
	(0.204)	(0.217)
neighBahçelievler_Kocasinan Merkez	-1.320^{***}	-0.905^{***}
	(0.201)	(0.214)
neighBahçelievler_Şirinevler	-1.343^{***}	-0.880^{***}
	(0.201)	(0.215)
neighBahçelievler_Siyavuşpaşa	-1.199^{***}	-0.742^{***}
	(0.200)	(0.214)
neighBahçelievler_Soğanlı	-1.349^{***}	-0.910^{***}
	(0.201)	(0.214)
neighBahçelievler_Yenibosna Merkez	-1.271^{***}	-0.854^{***}

	(0.200)	(0.214)
neighBahçelievler_Zafer	−1.455***	−0.989***
	(0.205)	(0.218)
neighBakırköy_Ataköy 1.	−0.260	0.346
	(0.244)	(0.255)
neighBakırköy_Ataköy 2-5-6.	0.080	0.608***
	(0.204)	(0.218)
neighBakırköy_Ataköy 3-4-11	−0.109	0.491**
	(0.212)	(0.226)
neighBakırköy_Ataköy 7-8-9-10	−0.318	0.217
	(0.201)	(0.216)
neighBakırköy_Basınköy	0.027	0.460**
	(0.210)	(0.224)
neighBakırköy_Cevizlik	−0.759***	−0.151
	(0.211)	(0.224)
neighBakırköy_Kartaltepe	−0.540***	0.053
	(0.201)	(0.216)
neighBakırköy_Osmaniye	−0.655***	−0.168
	(0.206)	(0.218)
neighBakırköy_Sakızağacı	−0.434**	0.158
	(0.208)	(0.221)
neighBakırköy_Şenlikköy	−0.042	0.436**
	(0.201)	(0.214)
neighBakırköy_Yenimahalle	−0.889***	−0.293
	(0.217)	(0.229)
neighBakırköy_Yeşilköy	0.055	0.626***
	(0.201)	(0.215)
neighBakırköy_Yeşilyurt	0.310	0.920***
	(0.206)	(0.222)
neighBakırköy_Zeytinlik	−0.714***	−0.116
	(0.220)	(0.233)
neighBakırköy_Zuhuratbaba	−0.546***	0.067
	(0.202)	(0.217)
neighBaşakşehir_Altınşehir	−1.295***	−1.014***
	(0.343)	(0.349)
neighBaşakşehir_Bahçeşehir 1. Kısım	−1.279***	−0.977***
	(0.201)	(0.215)
neighBaşakşehir_Bahçeşehir 2. Kısım	−1.574***	−1.189***
	(0.201)	(0.214)
neighBaşakşehir_Başak	−1.162***	−0.775***
	(0.204)	(0.217)
neighBaşakşehir_Başakşehir	−1.208***	−0.811***
	(0.203)	(0.217)
neighBaşakşehir_Güvercintepe	−1.898***	−1.515***
	(0.207)	(0.217)
neighBaşakşehir_İkitelli OSB	−1.084***	−0.695***
	(0.244)	(0.254)
neighBaşakşehir_Kayabaşı	−1.505***	−1.191***
	(0.203)	(0.214)
neighBaşakşehir_Şahintepe	−2.210***	−1.957***
	(0.343)	(0.348)
neighBaşakşehir_Ziya Gökalp	−1.319***	−0.961***
	(0.212)	(0.225)
neighBayrampaşa_Altıntepsi	−1.247***	−0.766***
	(0.206)	(0.219)
neighBayrampaşa_Cevatpaşa	−1.124***	−0.667*
	(0.343)	(0.350)
neighBayrampaşa_İsmet Paşa	−1.108***	−0.612**
	(0.244)	(0.254)
neighBayrampaşa_Kartaltepe	−1.219***	−0.699***
	(0.230)	(0.243)
neighBayrampaşa_Kocatepe	−0.984***	−0.511**
	(0.235)	(0.248)
neighBayrampaşa_Muratpaşa	−1.244***	−0.782***
	(0.212)	(0.225)
neighBayrampaşa_Ortamahalle	−1.232***	−0.735***
	(0.220)	(0.231)
neighBayrampaşa_Terazidere	−1.204***	−0.738***
	(0.220)	(0.231)
neighBayrampaşa_Yenidoğan	−1.280***	−0.752***
	(0.210)	(0.222)
neighBayrampaşa_Yıldırım	−1.129***	−0.668***
	(0.215)	(0.228)
neighBeşiktaş_Abbasağa	−0.469**	0.040
	(0.205)	(0.218)
neighBeşiktaş_Akat	0.161	0.705***

	(0.203)	(0.218)
neighBeşiktaş_Arnavutköy	0.497**	1.008***
	(0.207)	(0.217)
neighBeşiktaş_Balmumcu	−0.016	0.527**
	(0.217)	(0.230)
neighBeşiktaş_Bebek	0.859***	1.373***
	(0.201)	(0.214)
neighBeşiktaş_Cihannüma	−0.387*	0.068
	(0.213)	(0.225)
neighBeşiktaş_Dikilitaş	−0.289	0.273
	(0.202)	(0.216)
neighBeşiktaş_Etiler	0.331*	0.889***
	(0.201)	(0.216)
neighBeşiktaş_Gayrettepe	−0.271	0.308
	(0.202)	(0.216)
neighBeşiktaş_Konaklar	−0.495**	0.091
	(0.204)	(0.220)
neighBeşiktaş_Kültür	0.350	0.835***
	(0.226)	(0.237)
neighBeşiktaş_Kuruçeşme	0.363	0.826***
	(0.225)	(0.235)
neighBeşiktaş_Levazım	0.200	0.713***
	(0.203)	(0.218)
neighBeşiktaş_Levent	−0.045	0.495**
	(0.202)	(0.217)
neighBeşiktaş_Mecidiye	−0.413**	0.174
	(0.209)	(0.223)
neighBeşiktaş_Muradiye	−0.405**	0.114
	(0.206)	(0.219)
neighBeşiktaş_Nisbetiye	0.134	0.657***
	(0.203)	(0.218)
neighBeşiktaş_Ortaköy	−0.241	0.335
	(0.204)	(0.218)
neighBeşiktaş_Sinanpaşa	−0.279	0.175
	(0.243)	(0.254)
neighBeşiktaş_Türkali	−0.297	0.202
	(0.203)	(0.217)
neighBeşiktaş_Ulus	0.241	0.745***
	(0.201)	(0.215)
neighBeşiktaş_Vişnezade	0.186	0.692***
	(0.202)	(0.217)
neighBeşiktaş_Yıldız	−0.205	0.265
	(0.203)	(0.216)
neighBeykoz_Acarlar	−0.325	0.062
	(0.209)	(0.224)
neighBeykoz_Anadolu Hisarı	−0.369	0.202
	(0.243)	(0.255)
neighBeykoz_Kanlıca	0.033	0.584**
	(0.229)	(0.241)
neighBeykoz_Kavacık	−0.896***	−0.297
	(0.226)	(0.239)
neighBeykoz_Paşabahçe	−0.307	0.462
	(0.281)	(0.291)
neighBeykoz_Soğuksu	−0.456**	0.089
	(0.223)	(0.234)
neighBeykoz_Yalıköy	−0.379	0.291
	(0.256)	(0.263)
neighBeylikdüzü_Adnan Kahveci	−1.518***	−1.025***
	(0.200)	(0.214)
neighBeylikdüzü_Barış	−1.461***	−1.111***
	(0.202)	(0.216)
neighBeylikdüzü_Beylikdüzü OSB	−1.722***	−1.383***
	(0.201)	(0.215)
neighBeylikdüzü_Büyükşehir	−1.397***	−1.002***
	(0.205)	(0.218)
neighBeylikdüzü_Cumhuriyet	−1.771***	−1.322***
	(0.200)	(0.214)
neighBeylikdüzü_Dereağzı	−1.570***	−1.155***
	(0.223)	(0.234)
neighBeylikdüzü_Gürpınar	−1.608***	−1.140***
	(0.204)	(0.216)
neighBeylikdüzü_Kavaklı	−1.553***	−1.232***
	(0.204)	(0.217)
neighBeylikdüzü_Marmara	−1.400***	−0.974***
	(0.205)	(0.220)
neighBeylikdüzü_Sahil	−1.603***	−1.302***

	(0.281)	(0.289)
neighBeylikdüzü_Yakuplu	−1.776***	−1.313***
	(0.202)	(0.216)
neighBeyoğlu_Arap Cami	1.287***	1.834***
	(0.343)	(0.350)
neighBeyoğlu_Bereketzade	0.392	1.000***
	(0.281)	(0.290)
neighBeyoğlu_Camiikebir	−0.946***	−0.354
	(0.235)	(0.247)
neighBeyoğlu_Çatma Mescit	−0.390	0.212
	(0.256)	(0.268)
neighBeyoğlu_Cihangir	−0.067	0.491**
	(0.203)	(0.217)
neighBeyoğlu_Çukur	−0.889***	−0.335
	(0.281)	(0.289)
neighBeyoğlu_Firuzağa	−0.356*	0.204
	(0.215)	(0.229)
neighBeyoğlu_Gümüşsuyu	0.108	0.656***
	(0.205)	(0.218)
neighBeyoğlu_Hacıahmet	−1.289***	−0.708***
	(0.230)	(0.241)
neighBeyoğlu_Hacımimi	0.248	0.835***
	(0.281)	(0.291)
neighBeyoğlu_Halıcıoğlu	−1.077***	−0.472*
	(0.235)	(0.248)
neighBeyoğlu_Kadı Mehmet	−1.202***	−0.583**
	(0.215)	(0.229)
neighBeyoğlu_Kalyoncu Kulluğu	−0.263	0.305
	(0.257)	(0.267)
neighBeyoğlu_Kamer Hatun	−0.822**	−0.250
	(0.343)	(0.349)
neighBeyoğlu_Kaptanpaşa	−1.261***	−0.650***
	(0.208)	(0.222)
neighBeyoğlu_Katip Mustafa Çelebi	0.366	0.920***
	(0.256)	(0.267)
neighBeyoğlu_Keçeci Piri	−1.188***	−0.605**
	(0.235)	(0.247)
neighBeyoğlu_Kılıçali Paşa	0.126	0.691***
	(0.227)	(0.239)
neighBeyoğlu_Kocatepe	−0.454	0.078
	(0.281)	(0.289)
neighBeyoğlu_Küçük Piyale	−1.051***	−0.435*
	(0.214)	(0.228)
neighBeyoğlu_Kulaksız	−1.209***	−0.591**
	(0.220)	(0.233)
neighBeyoğlu_Kuloğlu	−0.602**	−0.047
	(0.235)	(0.247)
neighBeyoğlu_Müeyyedzade	−0.622*	−0.036
	(0.343)	(0.350)
neighBeyoğlu_Ömer Avni	0.190	0.741***
	(0.223)	(0.235)
neighBeyoğlu_Örnektepe	−1.061***	−0.481*
	(0.235)	(0.247)
neighBeyoğlu_Piri Paşa	−1.143***	−0.556**
	(0.211)	(0.225)
neighBeyoğlu_Pürtelaş Hasan Efendi	−0.023	0.532**
	(0.226)	(0.238)
neighBeyoğlu_Şahkulu	0.029	0.606**
	(0.235)	(0.249)
neighBeyoğlu_Sururi Mehmet Efendi	−1.210***	−0.624***
	(0.230)	(0.242)
neighBeyoğlu_Sütlüce	−1.105***	−0.504**
	(0.207)	(0.221)
neighBeyoğlu_Tomtom	−0.513**	0.055
	(0.235)	(0.249)
neighBeyoğlu_Yahya Kahya	−0.863***	−0.289
	(0.230)	(0.242)
neighBüyükçekmece_19.May	−1.529***	−1.049***
	(0.223)	(0.234)
neighBüyükçekmece_Alkent 2000	−0.858***	−0.482**
	(0.226)	(0.241)
neighBüyükçekmece_Atatürk	−1.335***	−0.844***
	(0.206)	(0.218)
neighBüyükçekmece_Çakmaklı	−1.660***	−1.507***
	(0.230)	(0.238)
neighBüyükçekmece_Celaliye	−2.424***	−2.113***

	(0.280)	(0.289)
neighBüyükçekmece_Cumhuriyet	-1.289^{***}	-0.795^{***}
	(0.220)	(0.233)
neighBüyükçekmece_Dızdarıye	-1.336^{***}	-0.865^{***}
	(0.243)	(0.252)
neighBüyükçekmece_Ekinoba	-1.478^{***}	-1.086^{***}
	(0.211)	(0.224)
neighBüyükçekmece_Fatih	-1.259^{***}	-0.732^{***}
	(0.207)	(0.220)
neighBüyükçekmece_Güzelce	-1.444^{***}	-0.940^{***}
	(0.343)	(0.351)
neighBüyükçekmece_Hürriyet	-1.929^{***}	-1.392^{***}
	(0.281)	(0.288)
neighBüyükçekmece_Kamiloba	-1.584^{***}	-1.114^{***}
	(0.226)	(0.240)
neighBüyükçekmece_Kumburgaz	-1.559^{***}	-1.137^{***}
	(0.220)	(0.241)
neighBüyükçekmece_Mimaroba	-1.409^{***}	-1.011^{***}
	(0.202)	(0.216)
neighBüyükçekmece_Mimarsinan	-1.270^{***}	-0.898^{***}
	(0.226)	(0.236)
neighBüyükçekmece_Murat Çeşme	-1.614^{***}	-1.136^{***}
	(0.211)	(0.225)
neighBüyükçekmece_Pınartepe	-1.132^{***}	-0.646^{***}
	(0.208)	(0.221)
neighBüyükçekmece_Sinanoba	-1.276^{***}	-0.953^{***}
	(0.205)	(0.218)
neighBüyükçekmece_Türkoba	-1.227^{***}	-0.328
	(0.343)	(0.352)
neighBüyükçekmece_Yenimahalle	-1.261^{***}	-0.767^{***}
	(0.281)	(0.289)
neighÇatalca_Kaleiçi	-1.686^{***}	0.275
	(0.343)	(0.384)
neighÇekmeköy_Alemdağ	-1.042^{***}	-0.527^{**}
	(0.256)	(0.266)
neighÇekmeköy_Aydınlar	-1.584^{***}	-0.987^{***}
	(0.223)	(0.234)
neighÇekmeköy_Çamlık	-1.502^{***}	-1.172^{***}
	(0.215)	(0.228)
neighÇekmeköy_Çatalmeşe	-1.698^{***}	-1.036^{***}
	(0.281)	(0.289)
neighÇekmeköy_Cumhuriyet	-1.623^{***}	-1.020^{***}
	(0.226)	(0.238)
neighÇekmeköy_Ekşioğlu	-1.449^{***}	-0.686^{*}
	(0.343)	(0.351)
neighÇekmeköy_Güngören	-1.362^{***}	-0.728^{**}
	(0.343)	(0.350)
neighÇekmeköy_Hamidiye	-1.148^{***}	-0.879^{***}
	(0.211)	(0.224)
neighÇekmeköy_Kirazlıdere	-1.481^{***}	-0.966^{***}
	(0.281)	(0.290)
neighÇekmeköy_Mehmet Akif	-1.436^{***}	-1.181^{***}
	(0.217)	(0.228)
neighÇekmeköy_Merkez	-1.138^{***}	-0.960^{***}
	(0.202)	(0.216)
neighÇekmeköy_Mimar Sinan	-1.276^{***}	-1.053^{***}
	(0.207)	(0.221)
neighÇekmeköy_Nişantepe	-1.537^{***}	-0.844^{***}
	(0.226)	(0.241)
neighÇekmeköy_Ömerli	-0.524	-0.588^{*}
	(0.343)	(0.357)
neighÇekmeköy_Soğukpınar	-1.583^{***}	-0.861^{***}
	(0.256)	(0.267)
neighÇekmeköy_Sultançiftliği	-1.361^{***}	-0.716^{***}
	(0.215)	(0.229)
neighÇekmeköy_Taşdelen	-1.412^{***}	-0.762^{***}
	(0.212)	(0.225)
neighEsenler_Birlik	-1.664^{***}	-1.174^{***}
	(0.343)	(0.348)
neighEsenler_Fatih	-1.525^{***}	-1.099^{***}
	(0.208)	(0.219)
neighEsenler_Fevzi Çakmak	-1.416^{***}	-1.000^{***}
	(0.217)	(0.229)
neighEsenler_Havaalanı	-1.380^{***}	-0.873^{***}
	(0.235)	(0.244)
neighEsenler_Kazım Karabekir	-1.459^{***}	-1.013^{***}

	(0.223)	(0.234)
neighEsenler_Kemer	-0.900^{***}	-0.441^{*}
	(0.235)	(0.245)
neighEsenler_Menderes	-1.393^{***}	-0.968^{***}
	(0.218)	(0.232)
neighEsenler_Mimar Sinan	-1.285^{***}	-0.896^{***}
	(0.243)	(0.255)
neighEsenler_Nine Hatun	-1.497^{***}	-1.040^{***}
	(0.223)	(0.235)
neighEsenler_Tuna	-1.590^{***}	-1.116^{***}
	(0.211)	(0.222)
neighEsenyurt_Akçaburgaz	-1.899^{***}	-1.489^{***}
	(0.226)	(0.237)
neighEsenyurt_Akevler	-1.916^{***}	-1.411^{***}
	(0.206)	(0.220)
neighEsenyurt_Akşemseddin	-1.531^{***}	-0.924^{***}
	(0.217)	(0.229)
neighEsenyurt_Ardıçlı	-1.899^{***}	-1.485^{***}
	(0.213)	(0.224)
neighEsenyurt_Aşık Veysel	-1.489^{***}	-1.172^{***}
	(0.207)	(0.220)
neighEsenyurt_Atatürk	-1.805^{***}	-1.394^{***}
	(0.244)	(0.254)
neighEsenyurt_Bağlarçeşme	-1.976^{***}	-1.604^{***}
	(0.203)	(0.215)
neighEsenyurt_Balıkyolu	-2.066^{***}	-1.631^{***}
	(0.343)	(0.348)
neighEsenyurt_Barbaros Hayrettin Paşa	-1.537^{***}	-1.077^{***}
	(0.205)	(0.218)
neighEsenyurt_Battalgazi	-1.938^{***}	-1.564^{***}
	(0.235)	(0.245)
neighEsenyurt_Cumhuriyet	-1.660^{***}	-1.233^{***}
	(0.202)	(0.217)
neighEsenyurt_Esenkent	-1.518^{***}	-1.164^{***}
	(0.212)	(0.225)
neighEsenyurt_Eşkinoz	-1.997^{***}	-1.439^{***}
	(0.217)	(0.230)
neighEsenyurt_Fatih	-1.971^{***}	-1.499^{***}
	(0.203)	(0.216)
neighEsenyurt_Gökevler	-1.761^{***}	-1.280^{***}
	(0.243)	(0.254)
neighEsenyurt_Güzelyurt	-1.785^{***}	-1.410^{***}
	(0.202)	(0.216)
neighEsenyurt_Hürriyet	-1.732^{***}	-1.321^{***}
	(0.281)	(0.288)
neighEsenyurt_İncirtepe	-1.859^{***}	-1.257^{***}
	(0.217)	(0.230)
neighEsenyurt_İnönü	-1.885^{***}	-1.301^{***}
	(0.202)	(0.217)
neighEsenyurt_İstiklal	-2.103^{***}	-1.724^{***}
	(0.211)	(0.221)
neighEsenyurt_Koza	-1.466^{***}	-1.162^{***}
	(0.208)	(0.220)
neighEsenyurt_Mehmet Akif Ersoy	-1.725^{***}	-1.333^{***}
	(0.223)	(0.235)
neighEsenyurt_Mehterçeşme	-1.874^{***}	-1.481^{***}
	(0.205)	(0.219)
neighEsenyurt_Mevlana	-1.595^{***}	-1.209^{***}
	(0.214)	(0.228)
neighEsenyurt_Namık Kemal	-1.584^{***}	-1.113^{***}
	(0.223)	(0.236)
neighEsenyurt_Necip Fazıl Kısakürek	-1.809^{***}	-1.429^{***}
	(0.343)	(0.348)
neighEsenyurt_Orhan Gazi	-1.951^{***}	-1.676^{***}
	(0.343)	(0.348)
neighEsenyurt_Örnek	-2.071^{***}	-1.731^{***}
	(0.206)	(0.218)
neighEsenyurt_Pınar	-1.974^{***}	-1.550^{***}
	(0.205)	(0.218)
neighEsenyurt_Piri Reis	-1.557^{***}	-1.141^{***}
	(0.235)	(0.247)
neighEsenyurt_Saadetdere	-2.075^{***}	-1.562^{***}
	(0.207)	(0.221)
neighEsenyurt_Şehitler	-2.092^{***}	-1.723^{***}
	(0.343)	(0.347)
neighEsenyurt_Selahaddin Eyyubi	-2.010^{***}	-1.763^{***}

	(0.281)	(0.287)
neighEsenyurt_Süleymaniye	-1.867^{***}	-1.447^{***}
	(0.343)	(0.348)
neighEsenyurt_Sultaniye	-1.916^{***}	-1.409^{***}
	(0.343)	(0.349)
neighEsenyurt_Talatpaşa	-1.961^{***}	-1.461^{***}
	(0.203)	(0.217)
neighEsenyurt_Turgut Özal	-1.380^{***}	-0.826^{***}
	(0.243)	(0.255)
neighEsenyurt_Yenikent	-1.880^{***}	-1.499^{***}
	(0.207)	(0.219)
neighEsenyurt_Yeşilkent	-1.865^{***}	-1.492^{***}
	(0.205)	(0.219)
neighEsenyurt_Yunus Emre	-2.030^{***}	-1.446^{***}
	(0.281)	(0.289)
neighEsenyurt_Zafer	-1.562^{***}	-1.102^{***}
	(0.218)	(0.232)
neighEyüp_Akşemsettin	-1.303^{***}	-0.849^{***}
	(0.205)	(0.222)
neighEyüp_Alibeyköy	-1.325^{***}	-0.825^{***}
	(0.207)	(0.222)
neighEyüp_Çırçır	-1.412^{***}	-0.965^{***}
	(0.200)	(0.217)
neighEyüp_Düğmeciler	-1.221^{***}	-0.670^{***}
	(0.223)	(0.235)
neighEyüp_Emniyettepe	-1.379^{***}	-0.788^{***}
	(0.256)	(0.268)
neighEyüp_Esentepe	-1.395^{***}	-0.733^{***}
	(0.205)	(0.222)
neighEyüp_Eyüp Merkez	-1.262^{***}	-0.626^{***}
	(0.218)	(0.233)
neighEyüp_Göktürk Merkez	-0.589^{***}	-0.394^{*}
	(0.200)	(0.216)
neighEyüp_Güzeltepe	-1.185^{***}	-0.797^{***}
	(0.204)	(0.217)
neighEyüp_İslambey	-1.299^{***}	-0.790^{***}
	(0.230)	(0.242)
neighEyüp_Karadolap	-1.346^{***}	-0.867^{***}
	(0.201)	(0.218)
neighEyüp_Mimarsinan	-0.699^{***}	-0.227
	(0.209)	(0.225)
neighEyüp_Nişanca	-1.143^{***}	-0.540^{**}
	(0.211)	(0.225)
neighEyüp_Rami Cuma	-1.185^{***}	-0.708^{**}
	(0.281)	(0.290)
neighEyüp_Rami Yeni	-1.509^{***}	-1.013^{***}
	(0.281)	(0.290)
neighEyüp_Sakarya	-1.654^{***}	-1.000^{***}
	(0.281)	(0.292)
neighEyüp_Silahtarağa	-1.205^{***}	-0.611^{***}
	(0.218)	(0.232)
neighEyüp_Topçular	-0.875^{***}	-0.334
	(0.227)	(0.240)
neighEyüp_Yeşilpınar	-1.341^{***}	-0.855^{***}
	(0.202)	(0.218)
neighFatih_Aksaray	-0.941^{***}	-0.370
	(0.213)	(0.226)
neighFatih_Akşemsettin	-0.731^{***}	-0.168
	(0.218)	(0.231)
neighFatih_Ali Kuşçu	-0.972^{***}	-0.397
	(0.282)	(0.291)
neighFatih_Atikali	-0.672^{**}	-0.144
	(0.281)	(0.290)
neighFatih_Ayvansaray	-1.104^{***}	-0.538^{**}
	(0.226)	(0.237)
neighFatih_Balat	-1.049^{***}	-0.525^{**}
	(0.243)	(0.255)
neighFatih_Beyazıt	-0.587^{*}	0.057
	(0.344)	(0.351)
neighFatih_Binbirdirek	0.527	1.163^{***}
	(0.343)	(0.349)
neighFatih_Cerrahpaşa	-0.927^{***}	-0.363
	(0.217)	(0.230)
neighFatih_Dervişali	-1.049^{***}	-0.538^{**}
	(0.257)	(0.267)
neighFatih_Haseki Sultan	-0.808^{***}	-0.270

	(0.208)	(0.222)
neighFatih_Hırka-i Şerif	-0.776^{**}	-0.216
	(0.343)	(0.349)
neighFatih_İskenderpaşa	-0.896^{***}	-0.308
	(0.235)	(0.247)
neighFatih_Karagümrük	-0.833^{***}	-0.303
	(0.243)	(0.255)
neighFatih_Kemal Paşa	-1.394^{***}	-0.826^{**}
	(0.343)	(0.350)
neighFatih_Kocamustafapaşa	-1.035^{***}	-0.512^{**}
	(0.207)	(0.221)
neighFatih_Mevlanakapı	-1.070^{***}	-0.571^{**}
	(0.210)	(0.223)
neighFatih_Molla Gürani	-0.729^{***}	-0.197
	(0.210)	(0.224)
neighFatih_Nişanca	-0.603^{*}	0.051
	(0.343)	(0.349)
neighFatih_Şehremini	-0.916^{***}	-0.393^{*}
	(0.201)	(0.215)
neighFatih_Seyyid Ömer	-0.967^{***}	-0.451^{**}
	(0.201)	(0.215)
neighFatih_Silivrikapı	-0.954^{***}	-0.432^{*}
	(0.209)	(0.222)
neighFatih_Sümbül Efendi	-1.110^{***}	-0.584^{**}
	(0.220)	(0.233)
neighFatih_Topkapı	-0.806^{***}	-0.265
	(0.226)	(0.239)
neighFatih_Yavuz Sultan Selim	-0.900^{***}	-0.357
	(0.226)	(0.237)
neighFatih_Yedikule	-0.907^{***}	-0.355
	(0.220)	(0.233)
neighGaziosmanpaşa_Bağlarbaşı	-1.391^{***}	-0.831^{***}
	(0.215)	(0.229)
neighGaziosmanpaşa_Barbaros Hayrettinpaşa	-1.528^{***}	-1.040^{***}
	(0.343)	(0.350)
neighGaziosmanpaşa_Fevzi Çakmak	-0.748^{**}	-0.166
	(0.343)	(0.350)
neighGaziosmanpaşa_Hürriyet	-1.163^{***}	-0.652^{*}
	(0.343)	(0.350)
neighGaziosmanpaşa_Karadeniz	-1.265^{***}	-0.810^{***}
	(0.213)	(0.227)
neighGaziosmanpaşa_Karayolları	-0.773^{***}	-0.251
	(0.210)	(0.224)
neighGaziosmanpaşa_Karlıtepe	-1.413^{***}	-0.911^{***}
	(0.243)	(0.255)
neighGaziosmanpaşa_Merkez	-1.448^{***}	-0.897^{***}
	(0.213)	(0.227)
neighGaziosmanpaşa_Mevlana	-1.592^{***}	-1.071^{***}
	(0.281)	(0.290)
neighGaziosmanpaşa_Pazariçi	-0.595^{**}	-0.044
	(0.281)	(0.291)
neighGaziosmanpaşa_Sarıgöl	-1.289^{***}	-0.782^{***}
	(0.223)	(0.235)
neighGaziosmanpaşa_Şemsipaşa	-1.461^{***}	-0.942^{***}
	(0.343)	(0.350)
neighGaziosmanpaşa_Yeni	-1.573^{***}	-1.006^{***}
	(0.343)	(0.350)
neighGaziosmanpaşa_Yıldıztabya	-1.381^{***}	-0.809^{***}
	(0.226)	(0.239)
neighGüngören_Abdurrahman Nafiz Gürman	-0.873^{***}	-0.464^{**}
	(0.207)	(0.220)
neighGüngören_Akıncılar	-1.248^{***}	-0.774^{***}
	(0.220)	(0.233)
neighGüngören_Gençosman	-1.342^{***}	-0.802^{***}
	(0.256)	(0.267)
neighGüngören_Güneştepe	-1.513^{***}	-0.975^{***}
	(0.281)	(0.290)
neighGüngören_Güven	-1.023^{***}	-0.562^{**}
	(0.235)	(0.247)
neighGüngören_Haznedar	-1.153^{***}	-0.642^{***}
	(0.213)	(0.227)
neighGüngören_Mareşal Çakmak	-1.295^{***}	-0.838^{***}
	(0.243)	(0.254)
neighGüngören_Mehmet Nesih Özmen	-0.912^{***}	-0.515^{**}
	(0.248)	(0.258)

neighGüngören_Merkez	-1.343^{***}	-0.832^{***}
	(0.215)	(0.228)
neighGüngören_Sanayi	-1.008^{***}	-0.445
	(0.343)	(0.350)
neighKadıköy_19.May	-0.706^{***}	-0.156
	(0.201)	(0.216)
neighKadıköy_Acıbadem	-0.563^{***}	-0.048
	(0.203)	(0.218)
neighKadıköy_Bostancı	-0.797^{***}	-0.331
	(0.200)	(0.215)
neighKadıköy_Caddebostan	-0.322	0.122
	(0.200)	(0.214)
neighKadıköy_Caferağa	-0.339^{*}	0.208
	(0.202)	(0.217)
neighKadıköy_Dumlupınar	-0.956^{***}	-0.431^{*}
	(0.207)	(0.220)
neighKadıköy_Eğitim	-0.774^{***}	-0.258
	(0.207)	(0.221)
neighKadıköy_Erenköy	-0.457^{**}	0.031
	(0.200)	(0.214)
neighKadıköy_Fenerbahçe	-0.157	0.329
	(0.200)	(0.215)
neighKadıköy_Feneryolu	-0.574^{***}	-0.038
	(0.200)	(0.215)
neighKadıköy_Fikirtepe	-0.968^{***}	-0.440^{**}
	(0.210)	(0.223)
neighKadıköy_Göztepe	-0.520^{***}	-0.0001
	(0.200)	(0.214)
neighKadıköy_Hasanpaşa	-0.909^{***}	-0.367
	(0.214)	(0.228)
neighKadıköy_Koşuyolu	-0.427^{*}	0.143
	(0.226)	(0.238)
neighKadıköy_Kozyatağı	-0.716^{***}	-0.165
	(0.200)	(0.216)
neighKadıköy_Merdivenköy	-0.808^{***}	-0.224
	(0.202)	(0.216)
neighKadıköy_Osmanağa	-0.548^{***}	-0.034
	(0.208)	(0.222)
neighKadıköy_Rasimpaşa	-0.657^{***}	-0.171
	(0.209)	(0.222)
neighKadıköy_Sahrayı Cedit	-0.727^{***}	-0.167
	(0.203)	(0.216)
neighKadıköy_Suadiye	-0.353^{*}	0.081
	(0.200)	(0.214)
neighKadıköy_Zühtüpaşa	-0.537^{***}	-0.018
	(0.205)	(0.219)
neighKağıthane_Çağlayan	-1.379^{***}	-0.866^{***}
	(0.202)	(0.216)
neighKağıthane_Çeliktepe	-1.023^{***}	-0.480^{**}
	(0.202)	(0.216)
neighKağıthane_Emniyetevleri	-0.969^{***}	-0.450^{**}
	(0.205)	(0.220)
neighKağıthane_Gültepe	-1.290^{***}	-0.665^{***}
	(0.203)	(0.217)
neighKağıthane_Gürsel	-1.215^{***}	-0.696^{***}
	(0.205)	(0.219)
neighKağıthane_Hamidiye	-1.133^{***}	-0.740^{***}
	(0.208)	(0.221)
neighKağıthane_Harmantepe	-1.238^{***}	-0.620^{***}
	(0.203)	(0.217)
neighKağıthane_Hürriyet	-1.302^{***}	-0.705^{***}
	(0.207)	(0.221)
neighKağıthane_Mehmet Akif Ersoy	-1.235^{***}	-0.652^{***}
	(0.218)	(0.232)
neighKağıthane_Merkez	-1.142^{***}	-0.702^{***}
	(0.202)	(0.216)
neighKağıthane_Nurtepe	-1.119^{***}	-0.612^{***}
	(0.208)	(0.222)
neighKağıthane_Ortabayır	-0.987^{***}	-0.421^{*}
	(0.205)	(0.218)
neighKağıthane_Sanayi	-1.233^{***}	-0.730^{***}
	(0.202)	(0.215)
neighKağıthane_Seyrantepe	-1.103^{***}	-0.686^{***}
	(0.205)	(0.217)
neighKağıthane_Şirintepe	-1.377^{***}	-0.835^{***}
	(0.209)	(0.221)

neighKağıthane_Talatpaşa	−1.253***	−0.728***
	(0.220)	(0.232)
neighKağıthane_Telsizler	−1.171***	−0.546**
	(0.206)	(0.219)
neighKağıthane_Yahyakemal	−1.265***	−0.750***
	(0.203)	(0.216)
neighKağıthane_Yeşilce	−1.210***	−0.765***
	(0.226)	(0.238)
neighKartal_Atalar	−1.137***	−0.622***
	(0.203)	(0.219)
neighKartal_Çarşı	−1.320***	−0.760***
	(0.213)	(0.227)
neighKartal_Çavuşoğlu	−1.183***	−0.728***
	(0.208)	(0.220)
neighKartal_Cevizli	−1.146***	−0.627***
	(0.204)	(0.219)
neighKartal_Cumhuriyet	−1.241***	−0.711***
	(0.208)	(0.222)
neighKartal_Esentepe	−1.285***	−0.843***
	(0.204)	(0.217)
neighKartal_Gümüşpınar	−1.351***	−0.877***
	(0.208)	(0.221)
neighKartal_Hürriyet	−1.463***	−0.921***
	(0.203)	(0.217)
neighKartal_Karlıktepe	−1.211***	−0.759***
	(0.203)	(0.216)
neighKartal_Kordonboyu	−0.835***	−0.330
	(0.203)	(0.216)
neighKartal_Orhantepe	−1.034***	−0.496**
	(0.210)	(0.224)
neighKartal_Orta	−1.213***	−0.715***
	(0.208)	(0.222)
neighKartal_Petroliş	−1.102***	−0.600***
	(0.203)	(0.218)
neighKartal_Soğanlık Yeni	−1.226***	−0.772***
	(0.204)	(0.219)
neighKartal_Topselvi	−1.348***	−0.818***
	(0.208)	(0.220)
neighKartal_Uğurmumcu	−1.399***	−0.934***
	(0.202)	(0.215)
neighKartal_Yakacık Yeni	−1.330***	−0.819***
	(0.207)	(0.220)
neighKartal_Yalı	−1.379***	−0.887***
	(0.210)	(0.224)
neighKartal_Yukarı	−1.149***	−0.661***
	(0.230)	(0.241)
neighKartal_Yunus	−1.328***	−0.766***
	(0.215)	(0.226)
neighKüçükçekmece_Atakent	−1.232***	−0.783***
	(0.200)	(0.215)
neighKüçükçekmece_Atatürk	−1.411***	−0.932***
	(0.205)	(0.219)
neighKüçükçekmece_Cennet	−1.241***	−0.783***
	(0.201)	(0.214)
neighKüçükçekmece_Cumhuriyet	−1.360***	−0.924***
	(0.202)	(0.215)
neighKüçükçekmece_Fatih	−1.061***	−0.723***
	(0.235)	(0.246)
neighKüçükçekmece_Fevzi Çakmak	−1.471***	−0.979***
	(0.230)	(0.242)
neighKüçükçekmece_Gültepe	−1.379***	−0.860***
	(0.203)	(0.217)
neighKüçükçekmece_Halkalı Merkez	−1.100***	−0.612***
	(0.204)	(0.218)
neighKüçükçekmece_İnönü	−1.462***	−0.984***
	(0.210)	(0.222)
neighKüçükçekmece_İstasyon	−1.466***	−0.959***
	(0.207)	(0.219)
neighKüçükçekmece_Kanarya	−1.709***	−1.174***
	(0.209)	(0.221)
neighKüçükçekmece_Kartaltepe	−1.250***	−0.768***
	(0.218)	(0.231)
neighKüçükçekmece_Kemalpaşa	−1.442***	−0.958***
	(0.226)	(0.239)
neighKüçükçekmece_Küçükçekmece İkitelli OSB	−1.073***	−0.720**

	(0.343)	(0.350)
neighKüçükçekmece_Mehmet Akif	-1.627^{***}	-1.078^{***}
	(0.211)	(0.224)
neighKüçükçekmece_Söğütlü Çeşme	-1.571^{***}	-1.009^{***}
	(0.206)	(0.219)
neighKüçükçekmece_Sultan Murat	-1.385^{***}	-0.888^{***}
	(0.205)	(0.219)
neighKüçükçekmece_Tevfik Bey	-1.168^{***}	-0.725^{***}
	(0.208)	(0.220)
neighKüçükçekmece_Yeni Mahalle	-1.305^{***}	-0.844^{***}
	(0.203)	(0.216)
neighKüçükçekmece_Yeşilova	-1.341^{***}	-0.842^{***}
	(0.202)	(0.216)
neighMaltepe_Altayçeşme	-1.083^{***}	-0.569^{***}
	(0.202)	(0.216)
neighMaltepe_Altıntepe	-1.024^{***}	-0.542^{**}
	(0.200)	(0.216)
neighMaltepe_Aydınevler	-1.172^{***}	-0.679^{***}
	(0.204)	(0.220)
neighMaltepe_Bağlarbaşı	-1.042^{***}	-0.486^{**}
	(0.202)	(0.216)
neighMaltepe_Başıbüyük	-0.825^{***}	-0.421^{*}
	(0.211)	(0.223)
neighMaltepe_Büyükbakkalköy	-0.769^{**}	-0.449
	(0.343)	(0.348)
neighMaltepe_Cevizli	-1.240^{***}	-0.754^{***}
	(0.201)	(0.216)
neighMaltepe_Çınar	-0.982^{***}	-0.467^{**}
	(0.202)	(0.218)
neighMaltepe_Esenkent	-1.142^{***}	-0.737^{***}
	(0.244)	(0.254)
neighMaltepe_Feyzullah	-0.869^{***}	-0.287
	(0.202)	(0.217)
neighMaltepe_Fındıklı	-1.478^{***}	-1.057^{***}
	(0.207)	(0.220)
neighMaltepe_Girne	-1.347^{***}	-0.856^{***}
	(0.212)	(0.225)
neighMaltepe_Gülsuyu	-1.311^{***}	-0.851^{***}
	(0.256)	(0.267)
neighMaltepe_İdealtepe	-1.016^{***}	-0.507^{**}
	(0.200)	(0.216)
neighMaltepe_Küçükyalı	-0.992^{***}	-0.472^{**}
	(0.200)	(0.216)
neighMaltepe_Yalı	-0.807^{***}	-0.226
	(0.204)	(0.217)
neighMaltepe_Zümrütevler	-1.260^{***}	-0.895^{***}
	(0.202)	(0.215)
neighPendik_Ahmet Yesevi	-1.569^{***}	-1.171^{***}
	(0.210)	(0.226)
neighPendik_Bahçelievler	-1.388^{***}	-0.854^{***}
	(0.226)	(0.238)
neighPendik_Çamçeşme	-1.422^{***}	-0.830^{***}
	(0.281)	(0.291)
neighPendik_Çamlık	-1.456^{***}	-0.830^{***}
	(0.209)	(0.225)
neighPendik_Çınardere	-1.308^{***}	-0.816^{***}
	(0.226)	(0.237)
neighPendik_Doğu	-1.167^{***}	-0.662^{***}
	(0.230)	(0.242)
neighPendik_Dumlupınar	-1.420^{***}	-0.948^{***}
	(0.211)	(0.224)
neighPendik_Esenler	-1.473^{***}	-0.887^{***}
	(0.217)	(0.230)
neighPendik_Esenyalı	-1.494^{***}	-1.150^{***}
	(0.235)	(0.250)
neighPendik_Fatih	-1.554^{***}	-1.212^{***}
	(0.230)	(0.245)
neighPendik_Fevzi Çakmak	-1.489^{***}	-0.925^{***}
	(0.220)	(0.234)
neighPendik_Güllü Bağlar	-1.527^{***}	-0.911^{***}
	(0.217)	(0.231)
neighPendik_Güzelyalı	-1.446^{***}	-1.008^{***}
	(0.226)	(0.240)
neighPendik_Harmandere	-1.246^{***}	-0.748^{***}
	(0.218)	(0.232)
neighPendik_Kavakpınar	-1.558^{***}	-1.016^{***}

	(0.230)	(0.242)
neighPendik_Kaynarca	-1.378^{***}	-0.841^{***}
	(0.217)	(0.232)
neighPendik_Kurtköy	-1.477^{***}	-0.972^{***}
	(0.204)	(0.218)
neighPendik_Orhangazi	-1.547^{***}	-1.119^{***}
	(0.226)	(0.240)
neighPendik_Sapan Bağları	-1.139^{***}	-0.574^{**}
	(0.211)	(0.223)
neighPendik_Şeyhli	-1.585^{***}	-1.035^{***}
	(0.220)	(0.235)
neighPendik_Sülüntepe	-1.506^{***}	-0.913^{***}
	(0.235)	(0.248)
neighPendik_Velibaba	-1.509^{***}	-0.979^{***}
	(0.209)	(0.223)
neighPendik_Yayalar	-1.470^{***}	-0.872^{***}
	(0.215)	(0.229)
neighPendik_Yeni Mahalle	-1.226^{***}	-0.664^{***}
	(0.210)	(0.223)
neighPendik_Yenişehir	-1.426^{***}	-0.914^{***}
	(0.202)	(0.217)
neighPendik_Yeşilbağlar	-1.267^{***}	-0.726^{***}
	(0.230)	(0.241)
neighSancaktepe_Abdurrahmangazi	-1.490^{***}	-1.056^{***}
	(0.205)	(0.217)
neighSancaktepe_Akpınar	-1.627^{***}	-1.255^{***}
	(0.235)	(0.247)
neighSancaktepe_Atatürk	-1.670^{***}	-1.146^{***}
	(0.223)	(0.234)
neighSancaktepe_Emek	-1.183^{***}	-0.875^{***}
	(0.211)	(0.223)
neighSancaktepe_Eyüp Sultan	-1.476^{***}	-1.067^{***}
	(0.226)	(0.237)
neighSancaktepe_Fatih	-1.657^{***}	-1.191^{***}
	(0.208)	(0.218)
neighSancaktepe_İnönü	-1.799^{***}	-1.426^{***}
	(0.281)	(0.291)
neighSancaktepe_Kemal Türkler	-1.533^{***}	-1.062^{***}
	(0.220)	(0.234)
neighSancaktepe_Meclis	-1.478^{***}	-1.087^{***}
	(0.256)	(0.266)
neighSancaktepe_Merve	-1.785^{***}	-1.191^{***}
	(0.235)	(0.243)
neighSancaktepe_Mevlana	-1.900^{***}	-1.264^{***}
	(0.223)	(0.230)
neighSancaktepe_Osmangazi	-1.834^{***}	-1.479^{***}
	(0.210)	(0.222)
neighSancaktepe_Safa	-1.795^{***}	-1.147^{***}
	(0.218)	(0.226)
neighSancaktepe_Sarıgazi	-1.623^{***}	-1.262^{***}
	(0.223)	(0.236)
neighSancaktepe_Veysel Karani	-1.554^{***}	-1.248^{***}
	(0.217)	(0.230)
neighSancaktepe_Yenidoğan	-1.699^{***}	-1.081^{***}
	(0.211)	(0.220)
neighSancaktepe_Yunus Emre	-1.798^{***}	-1.143^{***}
	(0.214)	(0.223)
neighSarıyer_Ayazağa	-0.746^{***}	-0.232
	(0.203)	(0.214)
neighSarıyer_Bahçeköy Merkez	-0.258	0.167
	(0.343)	(0.354)
neighSarıyer_Baltalimanı	0.673^{***}	1.248^{***}
	(0.230)	(0.241)
neighSarıyer_Büyükdere	-0.383^{*}	0.005
	(0.210)	(0.220)
neighSarıyer_Çamlıtepe	0.189	0.609^{*}
	(0.343)	(0.349)
neighSarıyer_Çayırbaşı	-0.462^{*}	-0.147
	(0.244)	(0.253)
neighSarıyer_Darüşşafaka	-0.366^{*}	0.111
	(0.206)	(0.219)
neighSarıyer_Demirci	-0.630^{*}	-0.078
	(0.343)	(0.350)
neighSarıyer_Emirgan	0.394^{*}	0.965^{***}
	(0.205)	(0.218)
neighSarıyer_Ferahevler	0.119	0.644^{***}

	(0.235)	(0.246)
neighSarıyer_Huzur	−0.599***	−0.234
	(0.205)	(0.216)
neighSarıyer_İstinye	0.168	0.841***
	(0.201)	(0.214)
neighSarıyer_Kemer	0.125	0.512
	(0.343)	(0.354)
neighSarıyer_Kireçburnu	−0.112	0.350
	(0.256)	(0.265)
neighSarıyer_Kumköy	−0.722**	−0.025
	(0.343)	(0.353)
neighSarıyer_Maden	−0.421**	0.097
	(0.202)	(0.218)
neighSarıyer_Maslak	−0.762***	−0.189
	(0.200)	(0.215)
neighSarıyer_Merkez	−0.411*	0.091
	(0.225)	(0.238)
neighSarıyer_Pınar	0.378*	0.958***
	(0.217)	(0.229)
neighSarıyer_Poligon	−0.186	0.378
	(0.235)	(0.246)
neighSarıyer_Reşitpaşa	0.015	0.626*
	(0.343)	(0.349)
neighSarıyer_Rumeli Hisarı	0.654***	1.254***
	(0.215)	(0.226)
neighSarıyer_Tarabya	0.236	0.756***
	(0.204)	(0.217)
neighSarıyer_Yeniköy	0.652***	1.384***
	(0.203)	(0.217)
neighSarıyer_Zekeriyaköy	−0.715***	−0.129
	(0.204)	(0.222)
neighŞile_Balibey	−1.315***	0.061
	(0.256)	(0.365)
neighŞile_Çavuş	−1.005***	0.397
	(0.226)	(0.338)
neighŞile_Hacı Kasım	−1.199***	0.193
	(0.343)	(0.427)
neighSilivri_Alibey	−1.705***	−0.489*
	(0.208)	(0.268)
neighSilivri_Balaban	−1.806***	−1.286***
	(0.343)	(0.408)
neighSilivri_Cumhuriyet	−2.011***	−0.828**
	(0.343)	(0.381)
neighSilivri_Fatih Silivri	−1.979***	−0.764**
	(0.281)	(0.327)
neighSilivri_Mimarsinan Silivri	−1.733***	−0.620**
	(0.235)	(0.295)
neighSilivri_Piri Mehmet Paşa	−1.706***	−0.513*
	(0.218)	(0.277)
neighSilivri_Selimpaşa Merkez	−1.648***	−1.247***
	(0.217)	(0.237)
neighSilivri_Yeni	−1.636***	−0.447*
	(0.210)	(0.269)
neighŞişli_19.May	−0.520**	0.045
	(0.205)	(0.219)
neighŞişli_Bozkurt	−0.644***	−0.102
	(0.201)	(0.215)
neighŞişli_Cumhuriyet	−0.501**	0.061
	(0.203)	(0.217)
neighŞişli_Duatepe	−0.640***	−0.092
	(0.203)	(0.217)
neighŞişli_Ergenekon	−0.498**	0.034
	(0.212)	(0.226)
neighŞişli_Esentepe	−0.301	0.295
	(0.205)	(0.220)
neighŞişli_Eskişehir	−0.944***	−0.404*
	(0.203)	(0.217)
neighŞişli_Feriköy	−0.776***	−0.234
	(0.201)	(0.215)
neighŞişli_Fulya	−0.424**	0.171
	(0.202)	(0.216)
neighŞişli_Gülbahar	−1.105***	−0.444**
	(0.205)	(0.218)
neighŞişli_Halaskargazi	−0.249	0.292
	(0.211)	(0.225)
neighŞişli_Halide Edip Adıvar	−0.762***	−0.220

	(0.206)	(0.220)
neighŞişli_Harbiye	0.273	0.795***
	(0.212)	(0.226)
neighŞişli_H. Rıfat Paşa	−0.561**	−0.069
	(0.243)	(0.255)
neighŞişli_İnönü	−0.608***	−0.082
	(0.226)	(0.238)
neighŞişli_İzzetpaşa	−1.182***	−0.559**
	(0.230)	(0.242)
neighŞişli_Kaptan Paşa	−0.392	0.116
	(0.256)	(0.267)
neighŞişli_Kuştepe	−1.106***	−0.451**
	(0.210)	(0.223)
neighŞişli_Mecidiyeköy	−0.821***	−0.184
	(0.202)	(0.215)
neighŞişli_Merkez	−0.499**	0.048
	(0.201)	(0.216)
neighŞişli_Meşrutiyet	−0.385*	0.174
	(0.207)	(0.221)
neighŞişli_Paşa	−0.980***	−0.445**
	(0.203)	(0.217)
neighŞişli_Teşvikiye	0.144	0.711***
	(0.202)	(0.216)
neighŞişli_Yayla	−1.150***	−0.642***
	(0.207)	(0.219)
neighSultanbeyli_Adil	−1.461***	−1.160***
	(0.204)	(0.215)
neighSultanbeyli_Ahmet Yesevi	−1.411***	−1.090***
	(0.343)	(0.349)
neighSultanbeyli_Battalgazi	−1.525***	−1.243***
	(0.343)	(0.348)
neighSultanbeyli_Hasanpaşa	−1.678***	−1.426***
	(0.235)	(0.246)
neighSultanbeyli_Mehmet Akif	−1.624***	−1.262***
	(0.281)	(0.291)
neighSultanbeyli_Mimar Sinan	−1.816***	−1.502***
	(0.243)	(0.253)
neighSultanbeyli_Necip Fazıl	−1.626***	−1.038***
	(0.220)	(0.233)
neighSultangazi_Cebeci	−1.543***	−1.230***
	(0.281)	(0.287)
neighSultangazi_Esentepe	−1.590***	−1.146***
	(0.235)	(0.245)
neighSultangazi_Eski Habipler	−1.403***	−1.300***
	(0.343)	(0.346)
neighSultangazi_Uğur Mumcu	−1.159***	−0.612*
	(0.343)	(0.351)
neighSultangazi_Yunus Emre	−1.573***	−1.175***
	(0.230)	(0.241)
neighTuzla_Aydınlı	−1.640***	−1.054***
	(0.205)	(0.220)
neighTuzla_Aydınlı - Birlik OSB	−0.893***	−0.260
	(0.343)	(0.347)
neighTuzla_Aydıntepe	−1.472***	−1.124***
	(0.209)	(0.226)
neighTuzla_Cami	−0.983***	−0.223
	(0.211)	(0.230)
neighTuzla_Evliya Çelebi	−1.311***	−0.797***
	(0.206)	(0.223)
neighTuzla_İçmeler	−1.482***	−1.065***
	(0.256)	(0.271)
neighTuzla_İstasyon	−1.163***	−0.337
	(0.205)	(0.226)
neighTuzla_Mescit	−1.878***	−1.524***
	(0.343)	(0.349)
neighTuzla_Mimar Sinan	−1.544***	−0.809***
	(0.218)	(0.231)
neighTuzla_Orta	−1.679***	−1.164***
	(0.230)	(0.240)
neighTuzla_Postane	−0.974***	−0.159
	(0.207)	(0.228)
neighTuzla_Şifa	−1.611***	−0.832***
	(0.230)	(0.242)
neighTuzla_Yayla	−1.316***	−0.557**
	(0.203)	(0.222)
neighÜmraniye_Altınşehir	−1.358***	−0.979***

	(0.235)	(0.247)
neighÜmraniye_Armağan Evler	-1.294^{***}	-0.831^{***}
	(0.204)	(0.220)
neighÜmraniye_Aşağı Dudullu	-1.168^{***}	-0.940^{***}
	(0.243)	(0.255)
neighÜmraniye_Atakent	-1.154^{***}	-0.687^{***}
	(0.230)	(0.244)
neighÜmraniye_Atatürk	-1.027^{***}	-0.634^{***}
	(0.218)	(0.233)
neighÜmraniye_Çakmak	-1.329^{***}	-0.928^{***}
	(0.203)	(0.218)
neighÜmraniye_Çamlık	-1.299^{***}	-0.859^{***}
	(0.210)	(0.223)
neighÜmraniye_Elmalıkent	-1.163^{***}	-0.707^{***}
	(0.243)	(0.258)
neighÜmraniye_Esenevler	-1.232^{***}	-0.759^{***}
	(0.212)	(0.227)
neighÜmraniye_Esenkent	-1.352^{***}	-1.088^{***}
	(0.244)	(0.254)
neighÜmraniye_Esenşehir	-1.266^{***}	-0.992^{***}
	(0.281)	(0.289)
neighÜmraniye_Fatih Sultan Mehmet	-0.944^{***}	-0.482^{*}
	(0.235)	(0.250)
neighÜmraniye_Huzur	-1.700^{***}	-1.529^{***}
	(0.343)	(0.350)
neighÜmraniye_Ihlamurkuyu	-0.886^{***}	-0.479
	(0.343)	(0.350)
neighÜmraniye_İnkılap	-1.175^{***}	-0.685^{***}
	(0.210)	(0.227)
neighÜmraniye_İstiklal	-1.199^{***}	-0.751^{***}
	(0.205)	(0.221)
neighÜmraniye_Madenler	-1.109^{***}	-0.846^{***}
	(0.226)	(0.238)
neighÜmraniye_Mehmet Akif	-1.179^{***}	-0.697^{***}
	(0.211)	(0.226)
neighÜmraniye_Namık Kemal	-1.280^{***}	-0.840^{***}
	(0.208)	(0.225)
neighÜmraniye_Necip Fazıl	-1.199^{***}	-0.896^{***}
	(0.214)	(0.227)
neighÜmraniye_Parseller	-1.069^{***}	-0.864^{***}
	(0.244)	(0.255)
neighÜmraniye_Saray	-1.013^{***}	-0.544^{**}
	(0.244)	(0.258)
neighÜmraniye_Şerifali	-1.227^{***}	-0.726^{***}
	(0.204)	(0.218)
neighÜmraniye_Site	-1.248^{***}	-0.829^{***}
	(0.211)	(0.224)
neighÜmraniye_Tantavi	-0.866^{***}	-0.479^{**}
	(0.209)	(0.223)
neighÜmraniye_Tatlısu	-1.103^{***}	-0.596^{***}
	(0.208)	(0.222)
neighÜmraniye_Yaman Evler	-1.195^{***}	-0.715^{**}
	(0.281)	(0.294)
neighÜsküdar_Acıbadem	-0.483^{**}	0.070
	(0.203)	(0.218)
neighÜsküdar_Ahmediye	-0.770^{***}	-0.185
	(0.204)	(0.217)
neighÜsküdar_Altunizade	-0.604^{***}	-0.021
	(0.211)	(0.226)
neighÜsküdar_Aziz Mahmut Hüdayi	-0.592^{***}	-0.035
	(0.204)	(0.218)
neighÜsküdar_Barbaros	-0.408^{*}	0.161
	(0.223)	(0.235)
neighÜsküdar_Beylerbeyi	0.339	0.777^{***}
	(0.215)	(0.228)
neighÜsküdar_Bulgurlu	-1.103^{***}	-0.582^{***}
	(0.207)	(0.222)
neighÜsküdar_Burhaniye	0.044	0.586^{**}
	(0.281)	(0.291)
neighÜsküdar_Çengelköy	-0.285	0.327
	(0.209)	(0.224)
neighÜsküdar_Cumhuriyet	-1.145^{***}	-0.600^{***}
	(0.205)	(0.220)

neighÜsküdar_Ferah			-1.120^{***}	-0.683^{***}
			(0.226)	(0.239)
neighÜsküdar_Güzeltepe			-0.753^{**}	-0.180
			(0.343)	(0.351)
neighÜsküdar_İcadiye			-0.695^{***}	-0.096
			(0.202)	(0.217)
neighÜsküdar_Kandilli			-0.178	0.451^{**}
			(0.210)	(0.222)
neighÜsküdar_Kısıklı			-0.908^{***}	-0.378
			(0.235)	(0.249)
neighÜsküdar_Küçük Çamlıca			-0.517^{**}	-0.047
			(0.244)	(0.256)
neighÜsküdar_Küçüksu			-0.091	0.682^{**}
			(0.281)	(0.291)
neighÜsküdar_Kuleli			1.116^{***}	1.765^{***}
			(0.281)	(0.288)
neighÜsküdar_Kuzguncuk			0.560^{**}	1.074^{***}
			(0.256)	(0.267)
neighÜsküdar_Mimar Sinan			-0.867^{***}	-0.257
			(0.202)	(0.217)
neighÜsküdar_Murat Reis			-0.894^{***}	-0.307
			(0.203)	(0.217)
neighÜsküdar_Salacak			-0.364^{*}	0.178
			(0.203)	(0.217)
neighÜsküdar_Selamiali			-0.829^{***}	-0.223
			(0.203)	(0.217)
neighÜsküdar_Selimiye			-0.655^{***}	-0.106
			(0.213)	(0.225)
neighÜsküdar_Sultantepe			-0.465^{**}	0.106
			(0.203)	(0.218)
neighÜsküdar_Ünalan			-0.715^{***}	-0.191
			(0.212)	(0.225)
neighÜsküdar_Validei Atik			-0.921^{***}	-0.348
			(0.202)	(0.215)
neighÜsküdar_Yavuztürk			-0.659^{*}	-0.148
			(0.344)	(0.351)
neighÜsküdar_Zeynep Kamil			-0.914^{***}	-0.320
			(0.202)	(0.215)
neighZeytinburnu_Beştelsiz			-1.208^{***}	-0.677^{***}
			(0.244)	(0.254)
neighZeytinburnu_Çırpıcı			-1.135^{***}	-0.654^{***}
			(0.226)	(0.238)
neighZeytinburnu_Gökalp			-0.969^{***}	-0.391^{*}
			(0.212)	(0.224)
neighZeytinburnu_Kazlıçeşme			0.119	0.634^{***}
			(0.236)	(0.246)
neighZeytinburnu_Maltepe			-1.122^{***}	-0.564^{**}
			(0.210)	(0.224)
neighZeytinburnu_Merkezefendi			-0.579^{***}	-0.027
			(0.220)	(0.234)
neighZeytinburnu_Nuripaşa			-0.996^{***}	-0.493^{**}
			(0.205)	(0.217)
neighZeytinburnu_Seyitnizam			-0.880^{***}	-0.368
			(0.230)	(0.242)
neighZeytinburnu_Sümer			-0.765^{***}	-0.325
			(0.212)	(0.224)
neighZeytinburnu_Telsiz			-1.199^{***}	-0.615^{***}
			(0.223)	(0.235)
neighZeytinburnu_Veliefendi			-1.173^{***}	-0.682^{***}
			(0.218)	(0.230)
neighZeytinburnu_Yenidoğan			-1.113^{***}	-0.532^{**}
			(0.220)	(0.232)
neighZeytinburnu_Yeşiltepe			-1.188^{***}	-0.667^{***}
			(0.208)	(0.221)
Observations	25,219	25,219	25,219	25,219
R^2	0.830	0.840	0.896	0.898
Adjusted R^2	0.830	0.839	0.894	0.896
Residual Std. Error	0.353 (df = 25150)	0.343 (df = 25136)	0.279 (df = 24560)	0.276 (df = 24546)
F Statistic	$1{,}811.833^{***}$ (df = 68; 25150)	$1{,}607.943^{***}$ (df = 82; 25136)	323.078^{***} (df = 658; 24560)	322.996^{***} (df = 672; 24546)
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

APPENDIX C

ADDITIONAL TABLES FOR CHAPTER 3

Table C1. Point Moran's I for Independent Variables, Inference by Randomization. Weight Matrix is Defined as Unity Standardization of Inverse Distances within 30km Bands

	Point MI	Expected MI	Variance MI	Z Score	P Value	n
Price	0.1315	-0.0000	0.00000	139.60	0.00	25,219
Age	0.1507	-0.0000	0.00000	159.86	0.00	25,219
Area	0.0831	-0.0000	0.00000	88.13	0.00	25,219
# of Bathrooms	0.1057	-0.0000	0.00000	112.16	0.00	25,219
# of Bedrooms	0.0786	-0.0000	0.00000	83.42	0.00	25,219
# of Livingrooms	0.0475	-0.0000	0.00000	50.44	0.00	25,219
FloorTotal	0.2046	-0.0000	0.00000	217.09	0.00	25,219

Table C2. Moran's I of Regression Residuals, Inference by Normalization. Weight Matrix is Defined as Unity Standardization of Inverse Distances within 30 km Bands

	MI	Expected MI	Variance MI	Z Score	P Value	n
Model 1	0.1037	-0.0004	0.000000	114.63	0	25,219
Model 2	0.0691	-0.0005	0.000000	78.52	0	25,219
Δ(MI)	-0.0346			-38.14	0	
Model 3	0.03738	-0.0014	0.000000	44.52	0	25,219
Model 4	0.0295	-0.0015	0.000000	37.21	0	25,219
Δ(MI)	-0.0079			-9.04	0	

Table C3. Bivariate Moran's I of Regression Residuals and Relevant Independent Variables for All Models, Inference by Normalization, Weight Matrix is Defined as Unity Standardization of Inverse Distances within 30 km Bands

	Bivariate MI	Expected MI	Variance MI	Z Score	P Value	n
Model 1						
Resid.-Age	0.0294	0.0000	0.00000	62.01	0	25,219
Resid.-Area	0.0496	0.0000	0.00000	104.63	0	25,219
Resid.- floorTotal	-0.0107	0.0000	0.00000	-31.52	0	25,219
Model 2						
Resid.-Age	0.0190	0.0000	0.00000	40.02	0	25,219
Resid.-Area	0.0335	0.0000	0.00000	70.68	0	25,219
Resid.- floorTotal	-0.0130	0.0000	0.00000	-42.82	0	25,219
Model 3						
Resid.-Age	0.0092	0.0000	0.00000	19.38	0	25,219
Resid.-Area	0.0182	0.0000	0.00000	38.50	0	25,219
Resid.- floorTotal	0.0046	0.0000	0.00000	5.10	0	25,219
Model 4						
Resid.-Age	0.0115	0.0000	0.00000	24.24	0	25,219
Resid.-Area	0.0145	0.0000	0.00000	30.54	0	25,219
Resid.- floorTotal	0.0046	0.0000	0.00000	5.20	0	25,219

Table C4. Spatially Explaining the Residuals, All Models

	Dependent variable:			
	nWresid			
	Model 1	Model 2	Model 3	Model 4
Age_norm	0.011***	0.007***	0.004***	0.004***
	(0.001)	(0.001)	(0.0005)	(0.0004)
Area_norm	0.017***	0.009***	0.005***	0.004***
	(0.002)	(0.001)	(0.001)	(0.001)
Bathrooms_norm	0.011***	0.007***	0.005***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Livingrooms_norm	-0.001	0.0002	0.001**	-0.0001
	(0.001)	(0.001)	(0.001)	(0.0004)
Bedrooms_norm	-0.008***	−0.003***	-0.005***	-0.004***
	(0.002)	(0.001)	(0.001)	(0.001)
FloorTotal_norm	-0.005***	−0.005***	0.001	0.001***
	(0.001)	(0.001)	(0.0005)	(0.0004)
Observations	25,219	25,219	25,219	25,219
R^2	0.024	0.025	0.010	0.012
Adjusted R^2	0.023	0.024	0.010	0.012
Residual Std. Error	0.142	0.093	0.072	0.055
	(df = 25213)	(df = 25213)	(df = 25213)	(df = 25213)
F Statistic	102.124***	105.563***	42.369***	51.222***
	(df = 6; 25213)	(df = 6; 25213)	(df = 6; 25213)	(df = 6; 25213)
Note:	*p<0.1; **p<0.05; ***p<0.01			

APPENDIX D

FULL REGRESSION RESULTS – INITIAL MODEL, 6-DIGIT AND 7-DIGIT SAMPLES

	Dependent variable:	
	log(price)	
	6 digit sample	*7 digit sample*
Constant	12.841***	13.416***
	(0.113)	(0.124)
age	−0.001**	−0.004***
	(0.001)	(0.0005)
area	0.005***	0.004***
	(0.0003)	(0.0001)
bathrooms	0.072***	0.001
	(0.012)	(0.009)
living_rooms	−0.048**	−0.075***
	(0.021)	(0.016)
bedrooms	0.028***	−0.0001
	(0.011)	(0.008)
heating_typeKat Kaloriferi	−0.108	0.033
	(0.181)	(0.126)
heating_typeKlima	−0.451***	−0.161
	(0.133)	(0.150)
heating_typeKombi	−0.338***	−0.088
	(0.106)	(0.116)
heating_typeMerkezi	−0.286***	−0.070
	(0.107)	(0.117)
heating_typeMerkezi (Pay Ölçer)	−0.230**	−0.077
	(0.107)	(0.117)
heating_typeSoba	−0.548***	
	(0.133)	
property_typeResidence	0.103***	0.028
	(0.028)	(0.020)
storey_at_clusteredgroundLevel	−0.199***	−0.018
	(0.018)	(0.024)
storey_at_clusteredhalfStorey	−0.130***	−0.025
	(0.020)	(0.036)
storey_at_clusteredhighStorey	0.118***	−0.016
	(0.020)	(0.016)
storey_at_clusteredmiddleStorey	0.054***	−0.033**
	(0.014)	(0.015)

storey_at_clusteredtopLevel	-0.066^{***}	-0.073^{***}
	(0.018)	(0.017)
storey_at_clusteredundergroundLevel	-0.255^{***}	-0.094
	(0.038)	(0.069)
storey_at_clusteredveryHighStorey	0.159^{***}	0.107^{***}
	(0.032)	(0.023)
gatedHayır	-0.113^{***}	-0.029^{*}
	(0.018)	(0.015)
dist.2.closest.fire_department	0.00002^{***}	-0.00000
	(0.00001)	(0.00001)
dist.2.closest.hospital	-0.00001	0.00000
	(0.00001)	(0.00001)
dist.2.closest.shopping_mall	-0.00001	0.00002^{***}
	(0.00001)	(0.00001)
dist.2.closest.police_department	0.00001	0.00001
	(0.00001)	(0.00001)
dist.2.closest.transportation	0.00001	0.00003^{***}
	(0.00001)	(0.00001)
dist.2.closest.private.anaokulu	-0.0001^{***}	0.00000
	(0.00002)	(0.00002)
dist.2.closest.private.ilkokul	0.00003^{*}	0.0002^{***}
	(0.00002)	(0.00002)
dist.2.closest.private.lise	-0.00005^{***}	-0.0001^{***}
	(0.00001)	(0.00001)
dist.2.closest.private.ortaokul	-0.0001^{***}	-0.0001^{***}
	(0.00002)	(0.00002)
dist.2.closest.public.anaokulu	0.00000	-0.00004^{***}
	(0.00001)	(0.00001)
dist.2.closest.public.ilkokul	0.00003	-0.00001
	(0.00002)	(0.00002)
dist.2.closest.public.lise	-0.00002	0.00003^{**}
	(0.00002)	(0.00001)
dist.2.closest.public.ortaokul	0.00000	0.00004^{*}
	(0.00002)	(0.00002)
dist.2.closest.university	0.00003^{***}	0.0001^{***}
	(0.00001)	(0.00001)
districtAvcılar	-0.656^{***}	-0.790^{***}
	(0.044)	(0.143)
districtBağcılar	-0.438^{***}	-0.305^{***}
	(0.046)	(0.103)

districtBahçelievler	-0.272^{***}	0.055
	(0.037)	(0.041)
districtBakırköy	0.310^{***}	0.393^{***}
	(0.050)	(0.036)
districtBaşakşehir	-0.532^{***}	-0.690^{***}
	(0.047)	(0.059)
districtBayrampaşa	-0.148^{**}	-0.291
	(0.071)	(0.242)
districtBeşiktaş	0.390^{***}	0.402^{***}
	(0.050)	(0.034)
districtBeylikdüzü	-0.680^{***}	-0.565^{***}
	(0.037)	(0.081)
districtBeyoğlu	0.078	0.547^{***}
	(0.062)	(0.056)
districtBüyükçekmece	-0.515^{***}	-0.499^{***}
	(0.060)	(0.101)
districtÇekmeköy	-0.430^{***}	-0.223^{***}
	(0.056)	(0.063)
districtEsenler	-0.442^{***}	-0.212
	(0.067)	(0.173)
districtEsenyurt	-0.870^{***}	-1.017^{***}
	(0.038)	(0.149)
districtEyüp	-0.206^{***}	-0.557^{***}
	(0.044)	(0.063)
districtFatih	-0.090^{**}	0.272^{***}
	(0.046)	(0.097)
districtGaziosmanpaşa	-0.261^{***}	-0.156
	(0.072)	(0.096)
districtGüngören	-0.308^{***}	-0.175
	(0.079)	(0.180)
districtKadıköy	0.284^{***}	0.138^{***}
	(0.035)	(0.032)
districtKağıthane	-0.155^{***}	-0.028
	(0.039)	(0.092)
districtKartal	-0.278^{***}	-0.021
	(0.039)	(0.059)
districtKüçükçekmece	-0.311^{***}	-0.323^{***}
	(0.036)	(0.049)
districtMaltepe	-0.040	-0.159^{***}
	(0.036)	(0.040)

districtPendik	-0.415^{***}	-0.592^{**}
	(0.047)	(0.245)
districtSancaktepe	-0.622^{***}	-0.987^{***}
	(0.060)	(0.094)
districtSarıyer	0.252^{***}	0.172^{***}
	(0.050)	(0.037)
districtSilivri	-0.715^{***}	
	(0.184)	
districtŞişli	0.154^{***}	0.295^{***}
	(0.039)	(0.037)
districtSultanbeyli	-0.523^{***}	
	(0.086)	
districtTuzla	-0.333^{***}	0.203^{*}
	(0.066)	(0.107)
districtÜmraniye	-0.262^{***}	-0.215^{***}
	(0.045)	(0.060)
districtÜsküdar	0.078^{*}	0.315^{***}
	(0.040)	(0.040)
districtZeytinburnu	0.079	0.099
	(0.062)	(0.067)
Observations	3,088	4,217
R^2	0.745	0.485
Adjusted R^2	0.739	0.477
Residual Std. Error	0.249 (df = 3021)	0.239 (df = 4153)
F Statistic	133.448^{***} (df = 66; 3021)	62.054^{***} (df = 63; 4153)
Note:		*p<0.1; **p<0.05; ***p<0.01

APPENDIX E

ADDITIONAL TABLES FOR CHAPTER 4

Table E1. Checking the Effects of Truncation: Partial Results of Truncated Regression

	Dependent variable:	
	log(Price)	
	OLS	*truncated regression*
...	...	...
	...	...
digit_group1	−0.110***	−0.113***
	(0.007)	(0.006)
digit_group2	−0.076***	−0.077***
	(0.006)	(0.006)
digit_group3	−0.056***	−0.056***
	(0.006)	(0.006)
digit_group4	−0.035***	−0.037***
	(0.007)	(0.007)
digit_group5	−0.006	−0.008*
	(0.004)	(0.004)
digit_group6	0.014**	0.014**
	(0.006)	(0.006)
digit_group7	0.039***	0.040***
	(0.007)	(0.007)
digit_group8	0.066***	0.067***
	(0.006)	(0.006)
digit_group9	0.085***	0.089***
	(0.006)	(0.007)
as.factor(price_last6th)2	0.426***	0.423***
	(0.006)	(0.006)
as.factor(price_last6th)3	0.739***	0.735***
	(0.007)	(0.006)
as.factor(price_last6th)4	0.978***	0.977***
	(0.007)	(0.007)
as.factor(price_last6th)5	1.177***	1.173***
	(0.008)	(0.008)
as.factor(price_last6th)6	1.337***	1.334***
	(0.009)	(0.009)
as.factor(price_last6th)7	1.477***	1.472***
	(0.009)	(0.009)
as.factor(price_last6th)8	1.598***	1.594***
	(0.010)	(0.010)
as.factor(price_last6th)9	1.710***	1.764***
	(0.011)	(0.013)
Observations	3,088	
R^2	0.981	
Adjusted R^2	0.981	
Residual Std. Error	0.067 (df = 3003)	
F Statistic	1,878.699*** (df = 84; 3003)	
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table E2. Extending the Analysis to 7-digit Subsample: Explaining Residuals by Digit Group

	Dependent variable:	
	residuals	
	based on 5th digit	*based on 6th digit*
digit_group0	0.104***	0.069***
	(0.010)	(0.018)
digit_group1	−0.047	−0.057***
	(0.078)	(0.015)
digit_group2	−0.016	−0.020
	(0.034)	(0.017)
digit_group3	0.021	0.058***
	(0.038)	(0.018)
digit_group4	−0.057	0.057***
	(0.037)	(0.022)
digit_group5	0.075***	0.138***
	(0.010)	(0.016)
digit_group6	0.006	0.126***
	(0.033)	(0.021)
digit_group7	−0.003	0.151***
	(0.030)	(0.027)
digit_group8	0.017	0.187***
	(0.024)	(0.027)
digit_group9	0.016	0.189***
	(0.021)	(0.026)
Constant	−0.062***	−0.023***
	(0.008)	(0.004)
Observations	4,217	4,217
R^2	0.033	0.058
Adjusted R^2	0.031	0.055
Residual Std. Error (df = 4206)	0.234	0.231
F Statistic (df = 10; 4206)	14.419***	25.767***

Note: *p<0.1; **p<0.05; ***p<0.01

Table E3. Extending the Analysis to 7-digit Subsample: Residual Ratio Checks

	Dependent variable:	
	residualsRatios	
	based on 5th digit	*based on 6th digit*
digit_group0	0.007***	0.004***
	(0.001)	(0.001)
digit_group1	−0.003	−0.004***
	(0.005)	(0.001)
digit_group2	−0.001	−0.001
	(0.002)	(0.001)
digit_group3	0.002	0.004***
	(0.003)	(0.001)
digit_group4	−0.004	0.004***
	(0.003)	(0.001)
digit_group5	0.005***	0.009***
	(0.001)	(0.001)
digit_group6	0.0004	0.009***
	(0.002)	(0.001)
digit_group7	−0.0002	0.011***
	(0.002)	(0.002)
digit_group8	0.001	0.013***
	(0.002)	(0.002)
digit_group9	0.001	0.013***
	(0.001)	(0.002)
Constant	−0.005***	−0.002***
	(0.001)	(0.0003)
Observations	4,217	4,217
R^2	0.033	0.058
Adjusted R^2	0.031	0.056
Residual Std. Error (df = 4206)	0.016	0.016
F Statistic (df = 10; 4206)	14.313***	26.052***
Note:		*p<0.1; **p<0.05; ***p<0.01

Table E4. Extending the Analysis to Rentals: Explaining Residuals by Digit Group

	Dependent variable:
	residuals
	based on 3rd digit
digit_group0	0.106***
	(0.008)
digit_group1	0.003
	(0.010)
digit_group2	0.047***
	(0.009)
digit_group3	0.039***
	(0.009)
digit_group4	0.025**
	(0.011)
digit_group5	0.114***
	(0.006)
digit_group6	0.045***
	(0.009)
digit_group7	0.045***
	(0.010)
digit_group8	0.056***
	(0.009)
digit_group9	0.055***
	(0.010)
Constant	−0.036***
	(0.003)
Observations	15,673
R^2	0.029
Adjusted R^2	0.028
Residual Std. Error	0.244 (df = 15662)
F Statistic	46.706*** (df = 10; 15662)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table E5. Sign and Significance of Digit Groups in Different Models

	Dependent variable:					
	Residuals (7d)	Prices (7d)	Prices (7d)	Residuals (6d)	Prices (6d)	Prices (6d)
group by digit	6th	6th	6th	5th	5th	5th
			price levels controlled			*price levels controlled*
		...	...		...	...
digit_group_0	signf (+)	signf (+)	signf (-)	signf (+)	signf (-)	signf (-)
digit_group_1	signf (-)	signf (-)	signf (-)			signf (-)
digit_group_2		signf (-)	signf (+)		signf (-)	
digit_group_3	signf (+)	signf (+)	signf (-)		signf (-)	signf (-)
digit_group_4	signf (+)	signf (+)		signf (+)		signf (-)
digit_group_5	signf (+)	signf (+)	signf (+)	signf (+)	signf (+)	
digit_group_6	signf (+)	signf (+)	signf (+)	signf (+)		
digit_group_7	signf (+)	signf (+)	signf (+)	signf (+)		signf (+)
digit_group_8	signf (+)	signf (+)	signf (+)	signf (+)	signf (+)	signf (+)
digit_group_9	signf (+)	signf (+)	signf (+)	signf (+)	signf (+)	signf (+)
	...	...	...	...	...	...

APPENDIX F

ADDITIONAL FIGURES FOR CHAPTER 4

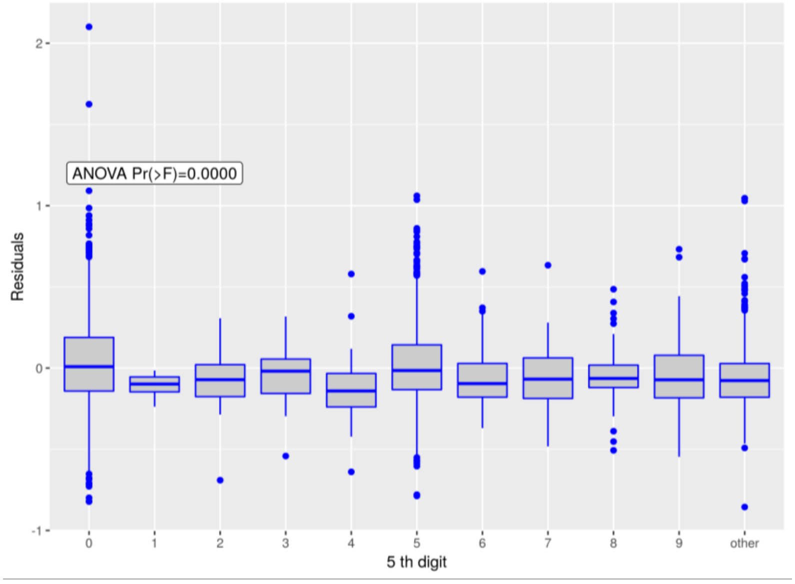

Figure F1 Pricing Equation Residuals by Price-Ending Groups and ANOVA, 7-digit Subsample, Grouping by 5th-digit

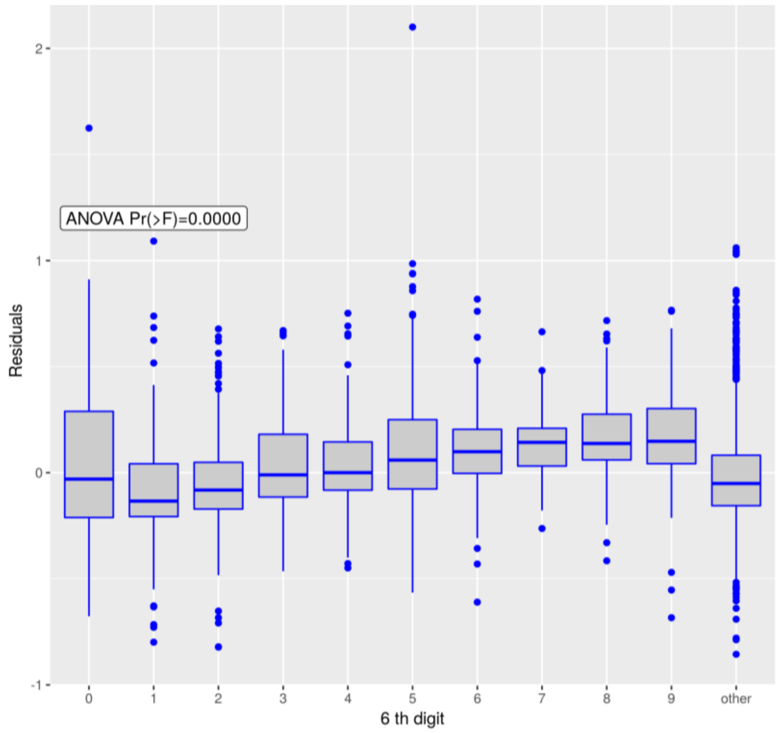

Figure F2 Pricing Equation Residuals by Price-Ending Groups and ANOVA, 7-digit Subsample, Grouping by 6th-digit

REFERENCES

Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Croz, J. D., Greenbaum, A., Hammerling, S., McKenney, A., & Sorensen, D. (1999). *LAPACK users' guide.* Third Edition. Philadelphia, PA: Society for industrial and applied mathematics.

Andrews, D., & Sánchez, A. C. (2011). The evolution of homeownership rates in selected OECD countries: Demographic and public policy influences. *OECD Journal: Economic Studies*, *2011*, 8–8. https://doi.org/10.1787/eco_studies-2011-5kg0vswqpmg2

Anselin, L. (1988). *Spatial econometrics: methods and models*. Dordrecht, The Netherlands: Springer. http://www.springer.com/us/book/9789024737352

Anselin, Luc. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, *27*(2), 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Anselin, Luc, Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*, *26*(1), 77–104. https://doi.org/10.1016/0166-0462(95)02111-6

Arbia, G. (2014). *A primer for spatial econometrics*. Palgrave Macmillan UK. //www.palgrave.com/us/book/9780230360389

Ball, C. A., Torous, W. N., & Tschoegl, A. E. (1985). The degree of price resolution: The case of the gold market. *Journal of Futures Markets*, *5*(1), 29–43. https://doi.org/10.1002/fut.3990050105

Basu, S., & Thibodeau, T. G. (1998). Analysis of spatial autocorrelation in house prices. *The Journal of Real Estate Finance and Economics*, *17*(1), 61–85. https://doi.org/10.1023/A:1007703229507

Baycan-Levent, T., & Gülümser, A. A. (2004). *Production and marketing of gated communities in Istanbul*. https://www.econstor.eu/handle/10419/117209

Bell, K. P., & Bockstael, N. E. (2000). Applying the generalized-moments estimation approach to spatial problems involving micro-level data. *The Review of Economics and Statistics, 82*(1), 72–82.

Beracha, E., & Seiler, M. J. (2014). The effect of listing price strategy on transaction selling prices. *The Journal of Real Estate Finance and Economics*, *49*(2), 237–255. https://doi.org/10.1007/s11146-013-9424-1

Bitter, C., Mulligan, G. F., & Dall'erba, S. (2007). Incorporating spatial variation in housing attribute prices: A comparison of geographically weighted regression and the spatial expansion method. *Journal of Geographical Systems*, *9*(1), 7–27. https://doi.org/10.1007/s10109-006-0028-7

Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2013). *Applied spatial data analysis with R* (2nd ed.). Springer-Verlag. https://www.springer.com/us/book/9781461476177

Bourassa, S., Cantoni, E., & Hoesli, M. (2010). Predicting house prices with spatial dependence: A comparison of alternative methods. *Journal of Real Estate Research*, *32*(2), 139–159. https://doi.org/10.5555/rees.32.2.l15423724383157x

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, *28*(4), 281–298. https://doi.org/10.1111/j.1538-4632.1996.tb00936.x

Burridge, P. (1980). On the cliff-ord test for spatial correlation. *Journal of the Royal Statistical Society. Series B (Methodological)*, *42*(1), 107–108.

Can, A. (1992). Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, *22*(3), 453–474. https://doi.org/10.1016/0166-0462(92)90039-4

Cardella, E., & Seiler, M. J. (2016). The effect of listing price strategy on real estate negotiations: An experimental study. *Journal of Economic Psychology*, *52*, 71–90. https://doi.org/10.1016/j.joep.2015.11.001

Chen, Y. (2015). A new methodology of spatial cross-correlation analysis. *PLOS ONE*, *10*(5), e0126158. https://doi.org/10.1371/journal.pone.0126158

Chun, Y., & Griffith, D. A. (2013). *Spatial statistics and geostatistics: Theory and applications for geographic information science and technology* (1st edition). SAGE Publications Ltd.

Cliff, A. D., & Ord, J. K. (1981). *Spatial processes: Models & applications*. Pion.

Cliff, A., & Ord, K. (1972). Testing for spatial autocorrelation among regression residuals. *Geographical Analysis*, *4*(3), 267–284. https://doi.org/10.1111/j.1538-4632.1972.tb00475.x

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates.

Cressie, N. A. C. (1993). *Statistics for spatial data*. J. Wiley.

Czaplewski, R. L., & Reich, R. M. (1993). *Expected value and variance of Moran's bivariate spatial autocorrelation statistic under permutation* [Research Paper RM-309]. U.S. Department of Agriculture, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO. http://hdl.handle.net/2027/umn.31951d02986843s

Dubin, R. A. (1988). Estimation of regression coefficients in the presence of spatially autocorrelated error terms. *The Review of Economics and Statistics*, *70*(3), 466–474. https://doi.org/10.2307/1926785

Dubin, R. A. (1998). Spatial autocorrelation: A primer. *Journal of Housing Economics*, *7*(4), 304–327. https://doi.org/10.1006/jhec.1998.0236

Dubin, R., Pace, R. K., & Thibodeau, T. G. (1999). Spatial autoregression techniques for real estate data. *Journal of Real Estate Literature*, *7*(1), 79–95.

Dunse, N., Jones, C., Orr, A., & Tarbet, H. (1998). The extent and limitations of local commercial property market data. *Journal of Property Valuation and Investment*, *16*(5), 455–473. https://doi.org/10.1108/14635789810237646

Florax, R., & Folmer, H. (1992). Specification and estimation of spatial linear regression models: Monte Carlo evaluation of pre-test estimators. *Regional Science and Urban Economics*, *22*(3), 405–432. https://doi.org/10.1016/0166-0462(92)90037-2

Geary, R. C. (1954). The contiguity ratio and statistical mapping. *The Incorporated Statistician*, *5*(3), 115. https://doi.org/10.2307/2986645

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis*, *24*(3), 189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

Giannetti, M., & Koskinen, Y. (2010). Investor protection, equity returns, and financial globalization. *Journal of Financial and Quantitative Analysis*, *45*(1), 135–168. https://doi.org/10.1017/S0022109009990524

Gillen, K., Thibodeau, T., & Wachter, S. (2001). Anisotropic autocorrelation in house prices. *The Journal of Real Estate Finance and Economics*, *23*(1), 5–30. https://doi.org/10.1023/A:1011140022948

Griffith, D. A. (2010). The Moran coefficient for non-normal data. *Journal of Statistical Planning and Inference*, *140*(11), 2980–2990. https://doi.org/10.1016/j.jspi.2010.03.045

Harris, L. (1991). Stock price clustering and discreteness. *The Review of Financial Studies*, *4*(3), 389–415. https://doi.org/10.1093/rfs/4.3.389

Huang, B., Wu, B., & Barry, M. (2010). Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science*, *24*(3), 383–401. https://doi.org/10.1080/13658810802672469

Jud, D., Seaks, T., & Winkler, D. (1996). Time on the market: The impact of residential brokerage. *Journal of Real Estate Research*, *12*(2), 447–458. https://doi.org/10.5555/rees.12.2.h617666h68168360

Kangallı Uyar, S. G., & Yayla, N. (2016). Konut fiyatlarinin hedonik fiyatlama yaklaşimina göre mekansal ekonometrik modeller ile tahmini. *Social Sciences (NWSASOS)*, *11*(4), 326–342.

Kaya, A. (2012). *Türkiye'de konut fiyatlarini etkileyen faktörlerin hedonik fiyat modeli ile belirlenmesi* [TCMB İstatistik Genel Müdürlüğü Uzmanlık Yeterlilik Tezi]. Türkiye Cumhuriyet Merkez Bankası.

Kelejian, H. H., & Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, *17*(1), 99–121. https://doi.org/10.1023/A:1007707430416

Kelejian, H. H., & Prucha, I. R. (2001). On the asymptotic distribution of the Moran I test statistic with applications. *Journal of Econometrics*, *104*(2), 219–257. https://doi.org/10.1016/S0304-4076(01)00064-1

Kelejian, H. H., & Prucha, I. R. (2007). HAC estimation in a spatial framework. *Journal of Econometrics*, *140*(1), 131–154. https://doi.org/10.1016/j.jeconom.2006.09.005

Kelejian, H. H., & Robinson, D. P. (1992). Spatial autocorrelation: A new computationally simple test with an application to per capita county police expenditures. *Regional Science and Urban Economics*, *22*(3), 317–331. https://doi.org/10.1016/0166-0462(92)90032-V

Keskin, B. (2008). Hedonic analysis of price in the Istanbul housing market. *International Journal of Strategic Property Management*, *12*(2), 125–138. https://doi.org/10.3846/1648-715X.2008.12.125-138

Keskin, B. (2010). *Alternative approaches to modelling housing market segmentation: Evidence from Istanbul*. http://etheses.whiterose.ac.uk/14983/1/531230.pdf

Knight, J. R. (2002). Listing price, time on market, and ultimate selling price: causes and effects of listing price changes. *Real Estate Economics*, *30*(2), 213–237. https://doi.org/10.1111/1540-6229.00038

Koramaz, T. K., & Dokmeci, V. (2012). Spatial determinants of housing price values in istanbul. *European Planning Studies*, *20*(7), 1221–1237. https://doi.org/10.1080/09654313.2012.673569

Kuo, W.-Y., Lin, T.-C., & Zhao, J. (2015). Cognitive limitation and investment performance: evidence from limit order clustering. *The Review of Financial Studies*, *28*(3), 838–875. https://doi.org/10.1093/rfs/hhu044

Lacetera, N., Pope, D. G., & Sydnor, J. R. (2012). Heuristic thinking and limited attention in the car market. *American Economic Review*, *102*(5), 2206–2236. https://doi.org/10.1257/aer.102.5.2206

Lesage, J. P. (1997). Bayesian estimation of spatial autoregressive models. *International Regional Science Review*, *20*(1–2), 113–129. https://doi.org/10.1177/016001769702000107

LeSage, J. P. (2000). Bayesian estimation of limited dependent variable spatial autoregressive models. *Geographical Analysis*, *32*(1), 19–35. https://doi.org/10.1111/j.1538-4632.2000.tb00413.x

LeSage, J. P., & Pace, R. K. (2007). A matrix exponential spatial specification. *Journal of Econometrics*, *140*(1), 190–214. https://doi.org/10.1016/j.jeconom.2006.09.007

LeSage, J. P., & Pace, R. K. (2009). *Introduction to spatial econometrics* (1 edition). Chapman and Hall/CRC.

LeSage, J. P., & Parent, O. (2007). Bayesian model averaging for spatial econometric models. *Geographical Analysis*, *39*(3), 241–267. https://doi.org/10.1111/j.1538-4632.2007.00703.x

Li, H., Calder, C. A., & Cressie, N. (2007). Beyond Moran's I: Testing for spatial dependence based on the spatial autoregressive model. *Geographical Analysis*, *39*(4), 357–375. https://doi.org/10.1111/j.1538-4632.2007.00708.x

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*(2 Part 1), 209–220.

Merlo, A., & Ortalo-Magné, F. (2004). Bargaining over residential real estate: Evidence from England. *Journal of Urban Economics*, *56*(2), 192–216. https://doi.org/10.1016/j.jue.2004.05.004

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, *37*(1/2), 17–23. https://doi.org/10.2307/2332142

Nguyen, A., Heeler, R. M., & Taran, Z. (2007). High-low context cultures and price-ending practices. *Journal of Product & Brand Management*, *16*(3), 206–214. https://doi.org/10.1108/10610420710751582

Ohnishi, T., Mizuno, T., Shimizu, C., & Watanabe, T. (2011). *On the evolution of the house price distribution*. https://doi.org/10.7916/D8794CJJ

Osland, L. (2010). An application of spatial econometrics in relation to hedonic house price modeling. *Journal of Real Estate Research*, *32*(3), 289–320. https://doi.org/10.5555/rees.32.3.d4713v80614728x1

Oust, A., Hansen, S. N., & Pettrem, T. R. (2019). Combining property price predictions from repeat sales and spatially enhanced hedonic regressions. *The Journal of Real Estate Finance and Economics*. https://doi.org/10.1007/s11146-019-09723-x

Ozus, E., Dokmeci, V., Kiroglu, G., & Egdemir, G. (2007). Spatial analysis of residential prices in istanbul. *European Planning Studies*, *15*(5), 707–721. https://doi.org/10.1080/09654310701214085

Pace, R. K., Barry, R., & Sirmans, C. F. (1998). Spatial statistics and real estate. *The Journal of Real Estate Finance and Economics*, *17*(1), 5–13. https://doi.org/10.1023/A:1007783811760

Palmon, O., A. Smith, B., & Sopranzetti, B. (2004). *Clustering in Real Estate Prices: Determinants and Consequences* (Vol. 26).

Reich, R. M., Czaplewski, R. L., & Bechtold, W. A. (1994). Spatial cross-correlation of undisturbed, natural shortleaf pine stands in northern Georgia. *Environmental and Ecological Statistics*, *1*(3), 201–217. https://doi.org/10.1007/BF00571392

Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology*, *7*(4), 532–547. https://doi.org/10.1016/0010-0285(75)90021-3

Sari, R., Ewing, B. T., & Aydin, B. (2007). Macroeconomic variables and the housing market in turkey. *Emerging Markets Finance and Trade*, *43*(5), 5–19. https://doi.org/10.2753/REE1540-496X430501

Sermaye Piyasası Kurulu, (2016). 2017 yılı gayrimenkul değerleme asgari ucret tarifesi. *Sermaye Piyasası Kurul Kararı No: 37/1314, Karar Tarihi: 29.12.2016.*

Sun, H., Tu, Y., & Yu, S.-M. (2005). A spatio-temporal autoregressive model for multi-unit residential market analysis. *The Journal of Real Estate Finance and Economics*, *31*(2), 155–187. https://doi.org/10.1007/s11146-005-1370-0

Tiefelsdorf, M., & Boots, B. (1995). The exact distribution of Moran's I. *Environment and Planning A: Economy and Space*, *27*(6), 985–999. https://doi.org/10.1068/a270985

Tiefelsdorf, Michael. (2002). The saddlepoint approximation of Moran's I's and local Moran's I's reference distributions and their numerical evaluation. *Geographical Analysis*, *34*(3), 187–206. https://doi.org/10.1111/j.1538-4632.2002.tb01084.x

Tu, Y., Sun, H., & Yu, S.-M. (2007). Spatial autocorrelations and urban housing market segmentation. *The Journal of Real Estate Finance and Economics*, *34*(3), 385–406. https://doi.org/10.1007/s11146-007-9015-0

Turkish Court of Accounts. (2008). *Büyükşehir belediyelerinde altyapi faaliyetlerinin koordinasyonu denetim raporu* [Performance Audit Report]. https://www.sayistay.gov.tr/tr/Upload/62643830/files/raporlar/diger/B%C3%BCy%C3%BCk%C5%9Fehir%20Belediyelerinde%20Altyap%C4%B1%20Faaliyetlerinin%20Koordinasyonu%20Performans%20Denetimi%20Raporu.pdf

Vega, S. H., & Elhorst, J. P. (2015). The Slx model. *Journal of Regional Science*, *55*(3), 339–363. https://doi.org/10.1111/jors.12188

Wartenberg, D. (1985). Multivariate spatial correlation: a method for exploratory geographical analysis. *Geographical Analysis*, *17*(4), 263–283. https://doi.org/10.1111/j.1538-4632.1985.tb00849.x

Wilhelmsson, M. (2002). Spatial models in real estate economics. *Housing, Theory and Society*, *19*(2), 92–101. https://doi.org/10.1080/140360902760385646

Yavas, A., & Yang, S. (1995). The strategic role of listing price in marketing real estate: theory and evidence. *Real Estate Economics*, *23*(3), 347–368. https://doi.org/10.1111/1540-6229.00668