The Effect of Competition on Residential Development Project Apartment Prices and Time on the Market

Lappeenranta–Lahti University of Technology LUT

Master's thesis in Strategic Finance and Analytics

2022

Sini Seppänen

Examiner: 1st Supervisor: Post-doctoral researcher Jan Stoklasa

2st Supervisor: Professor Pasi Luukka

ABSTRACT

Lappeenranta–Lahti University of Technology LUT

LUT School of Business and Management

Business Administration

Sini Seppänen

The Effect of Competition on Residential Development Project Apartment Prices and Time on the Market

Master's thesis

2022

96 pages, 27 figures, 18 tables and 8 appendices

Examiner: Post-doctoral researcher Jan Stoklasa and Professor Pasi Luukka

Keywords: Primary real estate markets, hedonic pricing, house prices, time-on-the-market, production of new housing

The purpose of this thesis is to examine one multi-phase residential development project located in Bratislava, Slovakia. The thesis aims to investigate how the competitive situation should be considered in the pricing of apartments. This is done by analysing which factors have an impact on the sale prices and time on the market (TOM) of apartments in residential development project. The analysis focuses on the primary residential real estate market where the seller is a construction company.

Hedonic pricing models are estimated using sale transaction data from residential development project between 2016 and 2021. The OLS technique is applied to explain apartment prices and TOM by using the apartment characteristics and market features as explanatory variables. The main interest is on the effects of competition, that is measured as the number of new apartments launched to sale by competitors. The results indicate that competition is significantly affecting TOM. Increased competition increases TOM for one- and two-room apartments but decreases for three- to four-room apartments. The competition has a positive impact on the prices of three- to four-room apartments, but there is no effect on other types of apartments. The findings reveal that also other factors affect house sale prices and TOM, and should be considered when pricing the apartments. Such as, interest rates have a significant effect on the price, and the launch season impacts TOM.

TIIVISTELMÄ

Lappeenrannan–Lahden teknillinen yliopisto LUT

LUT-kauppakorkeakoulu

Kauppatieteet

Sini Seppänen

Kilpailun vaikutukset uudiskohdehankkeen asuntojen hintoihin ja myyntiaikoihin

Kauppatieteiden pro gradu -tutkielma

2022

96 sivua, 27 kuvaa, 18 taulukkoa ja 8 liitettä

Tarkastaja: Tutkijatohtori Jan Stoklasa ja Professori Pasi Luukka

Avainsanat: Ensisijaiset asuntomarkkinat, hedoninen hinnoittelumalli, asuntojen hinnat, myyntiaika, uudiskohde

Tämän pro gradu -tutkielman tarkoituksena on tarkastella monivaiheista uudiskohteiden aluehanketta, joka sijaitsee Bratislavassa, Slovakiassa. Tutkimuksen tavoitteena on selvittää miten vallitseva kilpailutilanne tulisi huomioida asuntoja hinnoitellessa. Vastauksia pyritään löytämään selvittämällä mitkä tekijät vaikuttavat uudiskohteen asuntojen hintoihin ja myyntiaikoihin. Tämä tutkimus syventyy tutkimaan kilpailun vaikutuksia. Analyysi keskittyy ensisijaisiin asuntomarkkinoihin, jossa myyjänä toimii rakennusyhtiö.

Tutkimus analysoi aluehankkeen toteutuneita asuntokauppoja vuosilta 2016–2021 muodostamalla hedoniset hinnoittelumallit. Tutkimuksessa käytetään pienimmän neliösumman estimointimenetelmää selittämään asunnon hintaa ja myyntiaikaa käyttäen asunnon ominaisuuksia ja markkinamuuttujia selittävinä tekijöinä. Tulokset osoittavat, että kilpailulla on vaikutusta asuntojen myyntiaikoihin. Kasvanut kilpailu pidentää yksiöiden ja kaksioiden myyntiaikaa, mutta lyhentää kolmioiden. Kilpailulla on positiivinen vaikutus kolmioiden ja neliöiden toteutuneeseen myyntihintaan, mutta vaikutusta ei havaita yksiöiden ja kaksioiden kohdalla. Myös muut asunnon ominaisuudet ja markkinamuuttujat vaikuttavat asunnon hintoihin ja myyntiaikaan. Esimerkiksi korkotason kasvulla on merkittävä laskeva vaikutus asuntojen hintoihin, kun taas markkinoinnin aloitusajankohta vaikuttaa myyntiaikaan.

Table of contents

List of Figures

List of Tables

1. Introduction

This chapter introduces the topic and background of the thesis. The research questions are presented after which the data and research method used in this study are shortly discussed. Finally, the structure of the thesis is introduced.

1.1 Background of the study

Real estate market is having high importance in the economy. It plays an important role in the everyday life of individuals and households because of its consumption. Everyone needs a place to live and in addition, it works as an important investment function. (Łaszek, Olszewski, Waszczuk 2016, 189) The owner-occupied housing creates the greatest single source of wealth, and a house purchase is the largest investment decision for most of the people because it requires a lot of capital (DiPasquale & Wheaton 1992; Ooi & Thao 2012). Additionally, house price fluctuation has a significant impact on economic activity. Increasing house prices boosts economic activity when housing investments are growing and can cause a wealth effect that can be detected as increased consumption. (Cár 2006, 9)

Housing can be determined as a durable good. Therefore, new housing supply can be defined not only by the number of new units produced by the builder but also based on homeowner decisions to change the current housing stock. The new construction is the most significant addition to the housing stock. (DiPasquale 1999) A developer is most often a construction company that builds and sells newly built apartments to markets. The developer has conflicting incentives either to sell a house quickly at a lower price to achieve fast sales or to take higher prices and keep a house in the market for a longer period in order to maximize profits. (Zhou, Zahirovic-Herbert & Gibler 2017)

The first step for the seller of a house is to select the price the apartment is launched for sale. This price is often called a list price. Determining the optimal list price can be difficult and contains multiple decisions to make. If the house is mispriced, it can lower the final sales price. On the other hand, mispricing can also lead to unfavorable, such as too long marketing time. Time-on-the-market (TOM) is a measure of real estate market activity that can be

determined by supply and demand interactions. A high list price can keep the house on the market for a long period, indicating that pricing errors can lead to abnormal TOM. (Asabere & Huffman 1992)

Developers with large multi-phase residential development projects over a long period are dependent on long-term pricing strategies. This is mainly because the developer's actions have long-term consequences, and the decisions and efforts impact the performance of later phases. The price of a new housing unit should reflect the characteristics of the project valuation, and its decision-making process involves not only planning and positioning but also characteristics of surroundings (Ma, Liu & Sing 2020). The economic theory proposes that sellers are sensitive to the actions of other competitors in the market, and a price strategy should not be made without considering competitors' actions (Li 2004, 316). Consequently, the price of a new housing unit should reflect the existing competitive intensity (Ma et al. 2020). This thesis aims to study the effects of competition on developers' final sale prices and time-on-the-market.

The study provides an efficient way to achieve a comprehensive understanding of the effect of competition on residential real estate project sales by taking a deep look into one specific residential development project in Slovakia. To develop successful pricing strategies and price apartments correctly, it is important to understand what different factors impact apartment sales. In order to understand what factors can affect apartment sales, the thesis considers the house physical characteristics and selected market factors into the analysis.

This study has practical value because of many reasons. At first, no similar papers from Slovakia are found, making this the first study of its kind in Slovakia. Besides, According to Li (2004), previous studies on how a developer absorbs competitor behavior in its pricing strategy has been carried out marginally. A comprehensive body of literature on demand in real estate markets is extensively available, but supply is not so widely researched, mostly because of the lack of data (DiPasquale 1999; Łaszek et al. 2016, 192). Many studies focus only on aggregate data, and there is less data available where the unit of observation is the construction company. Some of the studies focus on the supply of single-family detached homes or renovation decisions but not the situations where the builder is involved. (DiPasquale 1999) Previous studies mainly focus on the secondary housing markets since it is the largest marketplace. However, the primary market is subject to many transactions but is researched much less. The primary housing markets consider the new construction,

whereas secondary markets the existing housing stock. (see e.g. Leszczynski & Olsewski 2017; Gustavsson & Vahtola 2014).

This study takes advantage of hedonic pricing models which have been commonly used in real estate market research. By creating hedonic regression models for project final sale prices and time-on-the-market using several different house characteristics and market factors, it can be observed which variables have an impact on the final sale prices of the project apartments. The methodology and variables used in this thesis are selected based on previous research. Sirmans, Macpherson, and Zietz (2005) introduce a comprehensive overview of real estate studies adopting hedonic pricing models and discuss the most used features used in these studies as explanatory variables. According to the authors, hedonic studies typically use single-stage equations, but a two-stage least squares model has also been used.

Most previous studies using the hedonic model framework make estimations using OLS. Using similar methods, these studies do not focus on competition effects on residential real estate project sales. However, many similar studies focus on either price or TOM, or both using the hedonic pricing models. Table 1 presents the most relevant studies in terms of methodology or the builder's perspective for this thesis. These studies adopt a hedonic pricing framework to study TOM and house prices or developer price-setting behavior. Studies on TOM are available less than studies on house prices. It is surprising when considering the importance of TOM for the seller.

Table 1: List of hedonic studies using OLS or 2SLS method

Author	**Dependent variable**	**Methodology**	**Main area of interest**
Barlindhaus & Nordahl (2017)	Price	OLS, log-linear	Developer's price setting behavior in residential development project
Wong, Chau, Yau & Cheung (2011)	Price	OLS, linear	Property price gradients
Li (2004)	Price, TOM	2SLS	Price and TOM trade-off in multiple-unit residential development projects
Yavas & Yang (1995b)	Price, TOM	2SLS	The strategic role of the listing price
Asabere & Huffman (1992)	TOM	OLS, log-linear	Macro-determinants of TOM
Kang & Kardner (1989)	Price, TOM	OLS, log-log	House prices and TOM in the residential real estate market

1.2 The objective of the study, research questions, and delimitation

This thesis focuses to examine apartment sale transactions in one multi-phase residential development project. The main focus of this thesis is to examine whether the competition has an impact on the project apartments' final sale prices and TOM. From the seller's perspective, it is important to understand the effects of competition, because as Li (2004) states, the pricing strategy should not be made without considering the competitor's actions. As many other factors may affect prices and TOM as well, the apartment characteristics, as well as selected market factors are included into the analysis. This is done to understand which factors have an impact on the final sale prices and TOM and at what intensities. This information should be considered when pricing the apartments of residential real estate development projects to reduce pricing errors. The seller should understand the impacts of the market conditions as well as the house characteristics to support the implementation of successful pricing strategies. The background of this thesis is to improve pricing strategies for a case company, that is a major construction company acting in several countries.

The research questions and main objectives for them are presented in Table 2. The first research question concerns suggesting a way to price apartments in a residential development project when considering the effects of the prevailing competitive situation. The conclusions and recommendation are given on setting the list price when considering the competitive situation. The final sale price is used as a proxy to make conclusions. The list price itself can not be used, because of missing features and information in the data. However, the list and final sale price are most often the same. One of the reasons for this is that the buyers usually do not have bargaining power. This means, the final sale price is suitable to use in this study to derive conclusions on the pricing. Two sub-questions are used to answer to the first research question. The first sub-question is about examining which factors have an impact on the final sale prices of the project. The second sub-question aims to find out which factor plays role in explaining the selling times in a project.

Table 2: Research questions and objectives

Main research question	**Objective**
1. How to price apartments of residential development project in different competitive market situations?	Reduce pricing errors and improve sales performance by pricing apartments correctly in different competitive situations.
Sub-questions to answer the main research question	
1.1 Which factors have an impact on the final sale prices of a residential development project?	Examine which variables impact the final sale price and find out does the competition plays a role in explaining the price.
1.2 Which factors explain the selling times (TOM) in a residential development project?	Examine which variables impact TOM and does the competition plays a role in explaining TOM.

The case project studied in this study is a multi-phase project located in Bratislava, Slovakia. More specifically, it is located in the Ružinov region, which belongs to Bratislava II borough. The project consists of multiple multi-story apartment buildings built in phases. Sales of apartments started in 2016. Prices in different regions in Slovakia differ from each other, thus it is reasonable to restrict the data to consider only competition from the nearby area (Cár 2006, 13). The thesis focuses on analyzing competitors and their projects in the Bratislava II (BA II) borough. The project is located also in this area, so it can be seen as a "nearby area", because most of the competitors in this area are located under four kilometer range. BA II covers the south-eastern part of Bratislava as shown in Figure 1.

Figure 1:Bratislava and its boroughs (Bratislava II 2021)

The analysis in this study focuses only on the primary residential real estate market in Bratislava, Slovakia and considers only the newly developed apartments, which can also be referred to as new housing production. This means, competition describes only the competition in the primary residential markets and excludes the competition from the secondary markets, which is the existing housing stock. This study refers to these apartments as housing units or apartments, and they are considered to be synonyms for each other. Marketing time is referred to as a time-on-the-market, selling time or TOM, each meaning the same.

1.3 Research method and data

A large developer company provides sales transaction data for this thesis. The data includes over 900 sale transactions from one residential real estate development project. The high-rise buildings were in different phases at the time data was gathered. The buildings are either in the planning, construction, or completed stage. The analysis focuses on examining nine phases, where each phase is a building containing 98 - 113 apartments.

The hedonic pricing model is often applied to investigate house prices using the special characteristics of the house (see e.g. Song & Quercia 2008; Wong et al. 2011). Aggregate models of the housing market have origins in the models of multi-sector macroeconomics that developed during the 1860s. The studies mainly focused on forecasting the new residential investments. Over time, price movements, owner-occupied housing, credit markets and the role of financial institutions began to arouse curiosity and interest. Finally, more advanced models were evolved. (DiPasquale & Wheaton 1994, 3)

This study uses hedonic pricing models to analyse the sale transaction data of one development project. As an apartment has a different characteristics that can all affect its value (Sirmans et al. 2005), this study uses these characteristics and selected market factors to estimate regression models separately for final sale prices and time-on-the-market. As Sirmans, MacDonald & Macpherson (2010) present, more than half of studies of TOM determine the hedonic model equations using the log-linear form, where the selling price is logged, and the TOM is unlogged. In contrast, 22 % of studies use both variables in logged form. Only 13 % use the linear form. Among these, even 83 % use the OLS estimation

method, whereas 17 % have adopted the two-stage least square method. This paper utilizes OLS regression which can be used to estimate coefficients in the model. OLS is a common way for fitting a model and it determines the parameters using minimization on the sum of squared errors (Studenmund 2010; Gareth et al. 2013, 21).

At first, the whole data set is used to estimate models for the final sale price and TOM. Additionally, separate models are constructed for each type of apartment, which means the one-room, two-room and three- to four-room apartments are studied separately. Different functional forms of each model are compared to find the best fit, and the best performing model results are presented. As a result, this study takes advantage of a linear, log-linear, and log-log functional forms, depending on the model. The analysis covers the period from the third quarter in 2016 to the second quarter in 2021.

Previous housing literature often measures the new supply by the number of new housing units build adjusted for demolition (Ooi & Thao 2011, 1445). Hui, Liang, Wang, and Wang (2016) suggest that one way to study the competition is to consider competitive intensity. As the aim of this thesis is to find out does the competition have an impact on the final sale price and TOM, this study measures the competitive intensity by the number of newly-built apartments in new projects nearby. To add, the study includes competitor price, overall supply, and interest rates of household mortgages into the analysis.

1.4 Structure of the thesis

This thesis is divided into six chapters. The first section focuses on briefly introducing the thesis, containing the research questions, objectives, limitations, method, and data used. The second chapter first introduces the special characteristics of the real estate markets. Then the Slovak housing markets and its history is presented, after which other characteristics of a real estate developer are introduced. The chapter continues with an introduction to the relevant literature and their main findings on house prices and time-on-the-market. The chapter ends with a section about previous studies on competition in housing markets.

Chapter 3 focuses on presenting the relevant theoretical framework for the study. The contribution of apartment characteristics is often determined using hedonic regression models (Sirmans et al. 2005). The chapter starts with an introduction to hedonic model

theory after which the model specification of hedonic models is discussed. The chapter also introduces five competitive forces presented by Porter (1989) to understand the nature of competition a company can face in residential real estate markets. Finally, chapter three takes a look at the search theory that is important to understand from the point of view of the time the apartment takes to sell. Search theory is widely used among TOM studies (Li 2004).

The fourth section starts with a short analysis of the project and its competitive environment, after which the variables used in this study are presented. The fifth section presents the results and reflects findings to previous studies. The last section concludes the thesis results with a discussion, presents the limitations, and gives suggestions for future research.

2 Background

This chapter gives a brief intro to the special characteristics of any housing. The project analyzed in this thesis is located in Slovakia thus, it is also relevant to understand the residential real estate market development and current market conditions in Slovakia. This chapter also explains the functioning of the primary residential real estate market from the developer's point of view. Furthermore, the previous related research and their main findings on house prices and time on the market are discussed.

2.1 Special characteristics of real estate markets

Housing units are unique in many ways. At first, people buy or rent a house to have a place to live, and housing is a necessity for households. There is asymmetric information in the markets since the seller and buyer do not share the same information of a house, for example, the seller has more information of the quality. Transaction costs can rise high because of multiple reasons, such as long searching times and broker costs. These kinds of characteristics can exist in other markets as well. The location factor is one of the special features of housing. Houses that are similar in physical features might be located in different areas so that they are valued differently. (Laakso, 1997, 24) Tobler (1979) states that the location is a key element in house price estimation.

The residential real estate market has plenty of other unique features that distinguish it from other conventional markets. Laakso (1997) describes housing as a multi-dimensional product, which comprises of characteristics of the house and building, but also as components associated with neighborhood and location or region. These characteristics are essential from the perspective of supply and demand. In the short run, the housing supply is relatively inelastic, and the number of new residential housing completed each year accounts for only 1-3 percent of the total stock. Existing housing stock has the greatest supply potential.

Oikarinen (2007) emphasizes heterogeneity, indivisibly and large size. No houses are precisely alike, and in addition to locational differences, they vary in view, construction

materials, shape, and age. These differences in housings make a house value estimation complex. The houses are usually large in size, and the investment requires most often a lot of capital. Housing can also be described as a durable good that provides a utility to its owners. This is why the pricing most likely differs from other assets. (Oikarinen 2007, 33, 37; Weimer & Hoyt 1984)

Actors and industries in the market, such as finance, construction, and construction materials as well as the economy have a consequential impact on the real estate industry. For example, the availability of house loans has had a significant effect on increasing house prices in Slovakia (Horvatova 2020, 1). The residential real estate market is a capital-intensive industry, making the barriers to entering the market remarkably greater than in many other markets. In addition to the high costs to start large projects, the knowledge and expertise need to be acquired. (Zhang 2010)

Real estate markets are vulnerable to economic changes and policies set by the government. Government actions can significantly impact real estate prices, such as through taxes, different fees policies, construction of public housing, or rental assistance, which raises the demand for housing services. The price response depends on the price elasticity of supply which determines the supply response. (DiPasquale 1999, 9; Zhang 2010)

An essential feature of residential real estate markets is that they are local. National real estate markets are a combination of regional residential real estate markets. The location constrains housing supply because the houses are not transferable. Similarly, the local households mostly generate the demand, even though migration changes demand between regions. (Laakso 1997)

Both the national and international conditions of the economy have an impact on the real estate markets, and the research at a national or regional level is a popular topic among the previous studies. The house's unique features and the effect of the economy on real estate markets highlight the importance of considering different factors in the analysis that can affect house prices (Laakso 1997).

2.2 Slovak real estate market overview

In order to understand the Slovak residential real estate market and its characteristics, a review of the market evolvement is presented in the following chapter. The chapter is followed by an overview of the house price development in Slovakia.

2.2.1 Development and current situation

The Slovak Republic housing policy has been developing for a long time while influenced by the geographical and political conditions, local economic, and the urbanization process. The country was centrally planned until 1989. The housing policy was also centralized and the state organized the housing for citizens. The country has made a lot of progress towards a market economy. Early signs of growth have been noticed since 1994. After 1989 many political and legislative changes took place, and the centrally controlled system changed into a market economy. Decreased inflation and budget deficit arose due to fiscal policies and prudent monetary. (Spirkova 2018, 21-22; United Nations 1999)

The Slovak Republic became independent in 1993. Housing policies went through multiple changes, and the state stepped aside from construction. The Slovak Rebubplik started to manage the frame documents and rules. They separated the housing policy into three different levels, local, central and regional. Between 1992-2006, a massive number of flats were privatized, leading to a decrease in the number of rental apartments. Overall, approximately 340 000 council flats and 250 000 cooperative flats were privatized. (Spirkova 2018, 22)

The banking sector was restructured in the early 2000s. After restructuring, the loans of the houses developed fast, and households indebted more than other CEE countries due to housing loans. As a result, by the end of 2020 house loans accounted one-third of banks' balance sheets and half of total loans issued. (Cesnak & Klacso 2021, 22) Slovakia's national bank provided the highest number of housing loans in the last quarter of 2020 that have been granted by then (Lexxus 2020b).

Slovakia's residential real estate markets play an important role in the economy as it has a high house ownership rate. The financial sector is dominated by banks and the country has

a tenuous capital market. One of the reasons housing stock is dominated by owner-occupied housing is low labour mobility. In 2020, most Slovaks lived in their own apartment, and the country faced a housing scarcity. The lack of housing dwellings is far behind the European average in terms of houses per 1000 inhabitants in 2011. (Liptáková 2020; Spirkova 2018, 21)

Ciarlone (2015) analysed the CEE countries' housing markets and found that the demand for houses is surprisingly high. Cesnak & Klacso (2021) states that the demand is not only high for the owner-occupied but also for investment purposes. Demand for houses is higher in areas with more job possibilities – which means larger cities like Bratislava. High demand is derived from the demographic development of the population. (Spirkova 2018, 21, 29) These facts emphasize the importance of primary housing market research in Slovakia. The new construction is the way to respond to increased demand (Ciarlone 2015).

Bratislava is the most efficient region in the economy of Slovak and also the most significantly growing housing market. According to Eurostat (2017), the Bratislava region's GDP per capita in purchasing power parity achieved 179 percent of the EU average, placing it eighth among the Union's regions in terms of economic performance. Bratislava is seen as a monocentric city where employment concentrates on the city center. Bratislava has continuously increasing demand for housing while the population has been growing since 1990. (Crowd Estate 2021; Rehák & Kácer 2018) The pace of new construction is not yet at the level that would meet rising demand due to the growing number of residents in Bratislava. (Liptáková 2021) Developers can not satisfy the demand needs with a rapid pace due to the lengthy process the development requires (Crowd Estate 2021). Covid 19 pandemic caused some delays in new construction projects in Slovakia, which delays the supply to meet the demand. Considering these, the construction sector of Slovakia has positive prospects in the long term. (European Comission 2021, 3)

The number of companies in the Slovak construction sector has slightly increased between 2010 and 2020. However, the industry faced a notable growth (+130 %) in activity in 2020. (European Commission 2021) Figure 2 presents how new house construction has evolved. It shows the number of houses under construction, new housing starts, and completed houses each year. The number of houses under construction rose unevenly from 1996 to 2020. The number of construction starts saw a peak in 2008, whereas the number of completed houses have been grown slowly and steadily during the period. Rybárova, Braunova, and Jantošová

(2016) point out that the construction industry in Slovakia reached positive results while Slovakia's accession to the EU during 2004-2008. This was mainly seen in a significant increase in the volume of construction production in 2008. The year 2012 has been described as a crisis for the construction industry when production decreased by 10 %. New housing starts saw their lowest level in 1996 and reached the top in 2008.

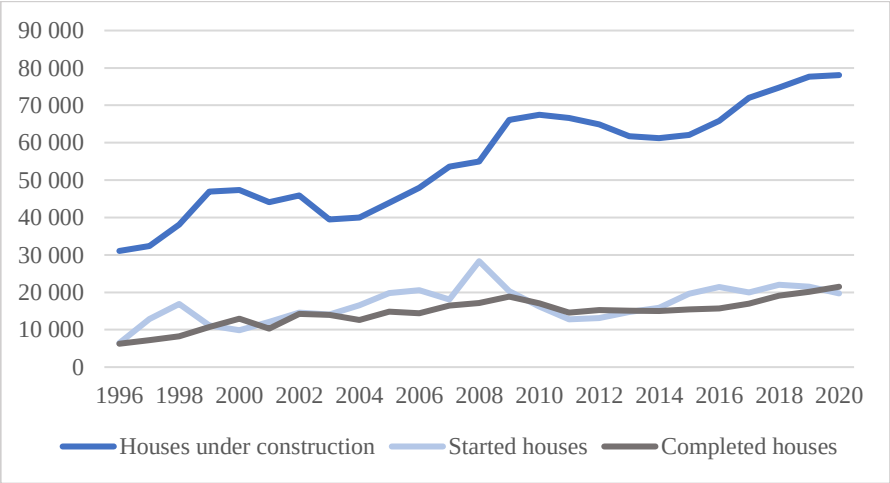


Figure 2: Started, completed, and under construction houses (Datacube 2021b)

Figure 3 describes the number of available and sold apartments as well as the average price of sold apartments in residential projects in each quarter in Bratislava during the years 2015 - 2021. On average, there have been approximately 2802 available apartments during the period. The number of available apartments reached its top in the second quarter of 2017 when there were 4304 available apartments. The number of available apartments began to decline after 2017 and reached its bottom in the first quarter of 2020. After the beginning of 2020, the number of available apartments has been quite steady. The lowest price 1666 €/m^2 is recorded in the first quarter of 2015.

Figure 3: Available and sold apartments and their average prices (without VAT) in Bratislava

Flats covers two-thirds of housing types in Slovakia. Besides, over 80 % of housing sales are flats in residential markets. (Spirkova 2018, 29) Figure 4 presents the structure of available apartments in Bratislava in new development projects in the first quarter of 2020. This proportion slightly differs among quarters and years, but traditionally, two-room apartments are available the most, followed by 3-room apartments. In contrast, five-room apartments are available much less.

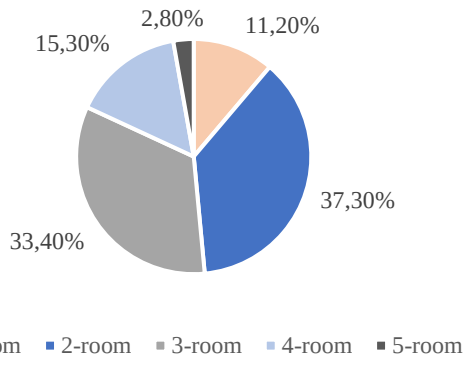


Figure 4: Structure of the available apartments in Bratislava (Lexxus 2020a)

2.2.2 Price development

Even though the Covid-19 pandemic negatively affected the economy by forcing to close businesses, the house prices in Slovakia have been rising. The house prices have increased

by 45,1 % from 2015 to 2021. The reasons for the price increase are the increased demand, low-interest rates, and growing population. (Liptáková 2021; European Commission 2021, 3) Interest rates on household mortgages have been decreasing since 2009, as shown in Figure 5. At the end of 2020, the interest rates for housing loans have been approximately 1 %. The interest rates for housing loans were very high, almost 8 % in the end of 2004. Low interest rates play a crucial role in housing financing (Horvatova 2020).

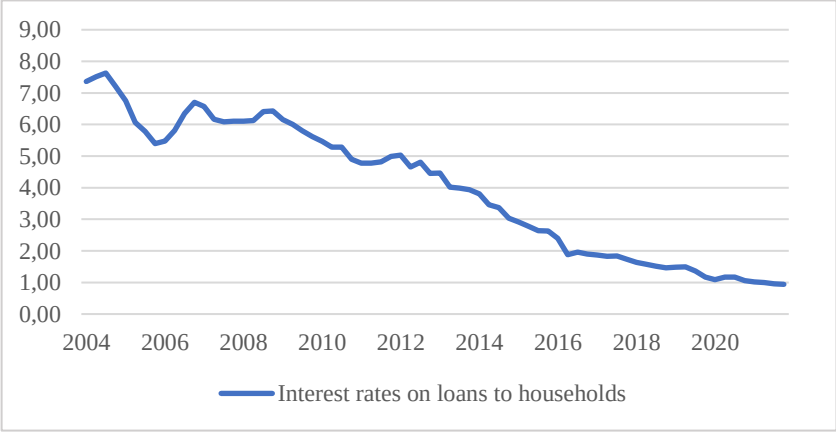


Figure 5: Interest rates on household mortgages (NBS 2022)

Figure 6 describes the house prices indices of new and existing dwellings overall in Slovakia from 2010 to 2021, where 2010 is used as the base year (Datacube 2021a). The prices have been rising quite steadily since 2014. According to Golej, Pánik, and Špirková (2016), the rapid price increase in residential real estate prices is a consequence of increased demand. Small drops in prices can be observed in 2017, 2018 and 2021. The prices hit bottom at the end of 2012. Further, existing apartment prices have been higher than new ones.

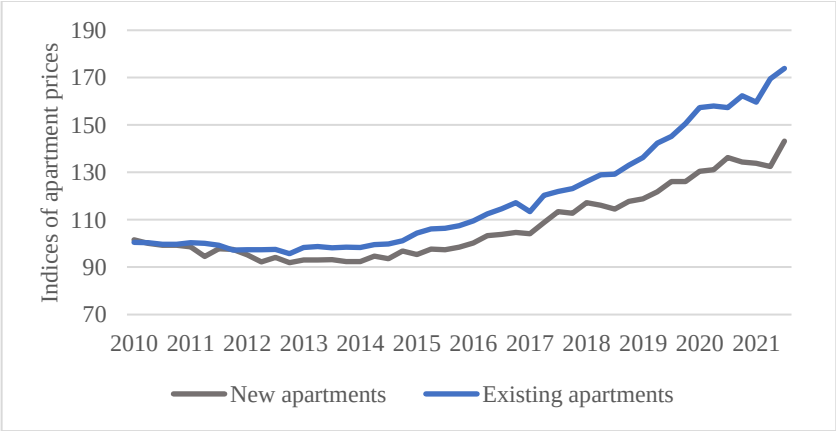


Figure 6: Price development of new and existing houses in Slovakia (2010 is used as a base year)

Considering the high demand for housing in Slovakia, especially in Bratislava, the construction sector of Slovakia has positive prospects in the long term. The prices have been growing quite rapidly and strong demand supports the house price growth also in the near future, although there are multiple other factors, such as interest rates that can affect house price development and prospects of the residential real estate industry.

2.3 Characteristics of real estate developers

The housing markets consist of two different segments, the primary and secondary markets. The primary market is where the houses are created which includes the developers' new apartments whereas the existing housing stock creates the secondary markets where the houses are sold by the homeowners. (Ooi & Thao 2012, 1437) The definition of a seller differs between these markets. The developers act as a seller in the primary markets whereas the households act as sellers and buyers in the secondary markets. (Laakso 1997, 24)

When the seller is a developer, it has its own characteristics compared to typical sellers selling a single house in a secondary market. A developer is most often a construction company that starts the process of constructing a new building and then selling newly built apartments to customers in order to make a profit. The builder has a limited amount of housing units to sell that are often launched into the market through presale. (Zhou et al. 2017, 191; Ho, Hui, Ho & Rengarajan 2013, 527) Development usually consists of multiple projects spread across years, and new units are gradually launched into the market (Barlindhaus & Nordahl 2017, 71). It should also be noted that different submarkets exist, including the business premises and administrative buildings (Laakso 1997, 24).

The primary residential real estate market has some interesting characteristics, such as they are characterized as oligopolistic. The prices are determined in an uncertain market and under incomplete information of competitors. In addition, the markets can be described as heterogeneous since the construction company can sell their apartments for different client segments at different prices. Besides, the length of the construction process distinguishes the market from other conventional markets. This means there is a particularly long time between the project's start and completion. The construction projects also require significant

investments and capital costs. Furthermore, one thing separating the market from other conventional markets is that the housing units are launched for sale before completion. (Ho et al. 2013, 527; Łaszek et al. 2016, 189; Zhou et al. 2017, 191)

The activities of developers usually cover the entire value chain of the real estate industry. Feasibility research, finance arrangements, project, design, construction, and property management are part of these activities. Moreover, the developer is taking care of the marketing and sales in most of the cases, and the whole process is relatively long. (Zhang 2010; Healey 1994) The construction process contains multi-stage decision-making processes. For example, the developer can choose to build a series of phases sequentially or develop the site at once. (Ott, Hughen & Read 2012)

New construction takes a big role in satisfying individual and household housing needs, although the new apartments cover only a smaller part of the total housing stock (Łaszek et al. 2016, 189). The new construction industry is highly volatile (Topel & Rosen 1988). Developers are usually defined as profit-seekers, and are more likely to start building more when the price levers are high. However, one can argue that developers fear intense price increases, leading to increased operating costs and weakening their competitiveness. When concerns of this exist, developers can start building more to press prices down. (Ooi & Thao 2012, 1437, 1440) Each of the developers in the markets are unique, and their actions and strategies vary from each other (Hui et al. 2016, 1204). However, presale practices are widely used among construction companies and are one thing that unites these companies. Next, this topic is discussed further.

2.3.1 Presale system

Nowadays, presale is the prevailing practice in real estate markets. The presale system is based on selling apartments before completing the construction. A risk that the demand will not meet expectations when a new building is completed is present for developers. Developer faces a financial risk when taking a loan and maintaining and holding a free apartment. The presale practice is widely used to lower the risks of the projects when the uncertainty of future demand exists. The method has gained tremendous popularity among developers,

especially in Europe, North America, and Asia. The method is advantageous, especially in large construction projects. (Lai, Ko & Yuqing 2004, 329-330)

The presale system allows a builder to evaluate the sales before committing investment funds. In the secondary housing market, the seller and buyer trade housing stock, that is the existing housing unit. Meanwhile, in primary housing markets, parties exchange the housing flow, which is the housing units that is not yet ready, but on the planning or construction stage. (Hua, Chang, Hsieh 2001, 81)

Lai et al. (2004) studied the characteristics of risk-and-return of the presale system. They suggest that the presale system is used for risk-sharing purposes and the developer can reduce the risk of bankruptcy by selling the housing units before finishing the construction. They modeled a presale decision using the real-options framework. Study results show that the presale system creates barriers for competitors to enter the market.

Many previous studies treat the presale system as a forward or futures contract (see e.g. Chang & Ward 1993). After the presale contract is signed, it is assumed that the buyer will eventually purchase the housing unit. There is a risk that the property value declines during the construction, and the buyer can default on the contract. Some contracts contain a clause allowing the buyer to terminate the agreement for a specific charge (Lai et al. 2004, 330-331). However, the contract commits the buyer to buy the housing unit. Most often the prices are fixed. (Barlindhaus & Nordahl 2017)

The sequence of events before the sale of a housing unit contains multiple steps. The sale process is described at the top level in Figure 7. Before starting the sales, the developer collects leads and opportunities, but any reservations can not be made at this point. The actual process of sales starts most often after having permission to construct. This means, after the building permit, the apartments are launched on the market. The listing price of a unit is usually a package price, which includes a storage and parking place. The agreements are concluded in two parts from which the final contract is made when finishing the construction.

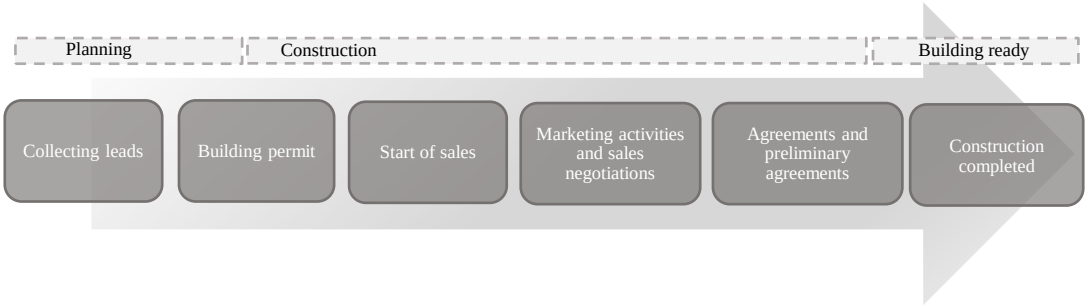


Figure 7: New development project sale process in Slovakia

Figure 8 presents well the special characteristics of the primary residential real estate market compared to other conventional markets. It describes the available completed and under-construction apartments in Bratislava. It can be easily observed that most of the apartments in the sale are under construction. Only a small fraction of apartments are completed. Most apartments will be sold during the presale phase, meaning the time before the unit is completed.

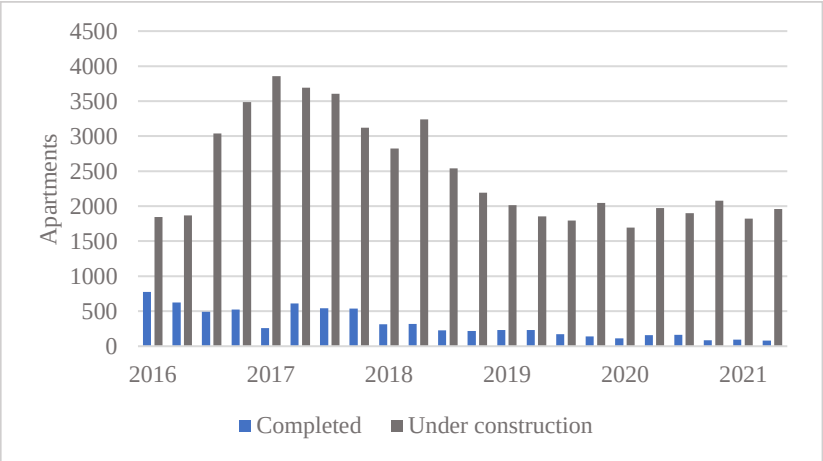


Figure 8: Available newly built apartments, completed vs. under construction

Demand uncertainty can be managed using other methods than presale as well. At first, the developer can wait and delay the construction process until the future looks brighter in terms of demand. This method contains an identification of a demand curve. (Lai et al. 2004, 329) For instance, Somerville (2001) emphasizes that a company should wait until demand conditions are clear even though the building permit has been granted. Titman (1985) determines this value as an “option to wait”. Likewise, Schwartz and Torous (2003) study results advocates for the “option to wait” theory.

When uncertainty about future demand prevails, the builder can lower the prices of housing units to ensure sufficient demand in the future. This method works best with projects that contain multiple phases. The builder can raise prices gradually if the future market situation looks promising. (Lai et al. 2004, 330) Sirmans, Turnbull, and Dombrow (1997) research results support this method. The authors studied the relationship between land price and residential development sequence. The results show that the developer starts by offering lower prices because uncertainty exists about a neighborhood's future characteristics. Over time, the prices begin to rise when more information about the neighborhood shows up.

2.3.2 Pricing Strategies

Construction companies are different in many dimensions. These unique features should be considered in the context of their pricing strategy. As Łaszek et al. (2016) argue, companies move along an aggregated demand curve, producing similar but different products to markets as developers compete to create a unique product. Because of this, developers might utilize bargaining strategies to find consumers' reservation points to maximize project profits.

In theory, builders adjust pricing to match the supply in the market (Hui et al. 2016). Sales comparison is a common tool for price setting. When a developer performs this kind of comparison, the different characteristics of a house should be considered to succeed in price setting within a project. Furthermore, the builder can deploy a hedonic pricing model to determine differences in prices between apartments and also, make a comparison between phases. The developer can select a site with the most complimentary amenities and neighborhood for the first stage. (Barlinhaus & Nordahl 2017; Pagourtzi, Assimakopoulos, Hartzixhristos, French 2003)

The pricing is not straightforward in large construction projects. The internal amenities of an area evolve as development progresses. The building progress can affect buyers' willingness to pay if they need to live on a construction site for a longer time. The developer can use a passive strategy to offer price discounts for the first phases of construction. This strategy has rising price levels when the completion date gets closer. Naturally, the prices in the last phase would have the highest prices.

The developer could also use sale pitches and take an offensive strategy to brand the new area by using artists' impressions. The company should be convincing that the area under development will be a high-quality residential area in the future. By doing intensive branding for the development, the builder wants to maximize the profits in the first phase. The prices in this kind of strategy will most likely remain flat over the development.

Different areas have different competitive intensities which should be reflected in pricing strategies. The competitive intensity can affect, for example, the profitability of selecting to add a premium. If one developer launches their apartments from the last phase on sale after which another project launches apartments from the first phase on sale, it can affect to premiums that the first developer is asking, if the second developer follows the strategy where they use lower prices in the first stage. (Barlinhaus & Nordahl 2017, 75-76)

Zhang (2010) arguments that different companies will face the competition differently since they adopt different strengths and resources. The competition strategy planning should start by recognizing the external environment, analysing the company's competitiveness, and considering all possible decision-making strategies.

Barlindhaus and Nordahl (2017) examined the developer's price setting behavior. They took the hedonic approach to analyse whether developers ask for lower prices on houses in redevelopment sites than in similar units in smaller developments. In addition, they research if there are differences in pricing in different construction phases. The findings show that developers with competitive redevelopment projects put higher prices on the last phases whereas redevelopment containing only one builder, the last phase has the lowest prices.

2.4 Relationship between prices and time-on-the-market

Previous research discusses about the trade-off between house price and the time it takes to sell. When a housing unit is put for sale, the seller selects an initial offer price. The offer price affects for the marketability of the unit. Setting a too high or too low price can negatively affect sales because it affects potential buyer's interest and willingness to buy. (Anglin, Rutherford, Springer 2003; Hui & Kung Yu 2012) This indicates the fact that the housing market is a search market where lower prices lead to faster sales (Gustavsson & Vahtola 2014).

Some previous research has studied the impact of listing price on TOM. Anglin et al. (2003) supports the hypothesis that a relationship exists between TOM and listing price. The authors present a model of the impact of pricing decisions of the seller using data of single-family housing sales from Texas. The findings indicate that there is no direct relationship between TOM and selling price. However, they vary along with the list price. The listing price affects the group of potential buyers thus, the higher the listing price, the longer the TOM will be. In addition, the effect of higher list price is magnified for houses in sale in a market with a low variance in the list prices.

Also, Knight (2002) argues that when the listing price is too high, the TOM gets longer, but also these apartments will eventually be sold for less because of mispricing. The author adopts the 2SLS method to examine this relationship. Similarly, Yavas & Yang (1995b) use both the 2SLS method and the search theory to study house prices and TOM. The authors suggest the list price only impacts TOM of mid-priced apartments and has no impact on low-priced or high-priced apartments. Their study mainly focuses on examining the relationship between TOM and the real estate agent. Cubbin (1974) found an inverse relationship between price and TOM. The study results indicates that buyers judge quality by price, so a seller faces difficulties in selling if the asking price is below average. The findings shows that the housing unit will be sold faster when the price is higher.

Most studies on time-on-the-market are based on investigating the relationship between price and TOM. When comparing hedonic pricing literature, most studies focus on examining price as the dependent variable, fewer part of them takes TOM as a dependent variable (Dubé & Legros 2015). Published research shows that there exists a robust relationship between prices and TOM across hedonic model specifications. A majority of the research on the relationship between TOM and price focuses on how TOM influences a house selling price. However, research seems to divide into two groups where one group supports the theory that the longer a house is on the market, the more likely it is to sell for a higher price. (Sirmans et al. 2010, 139-140; Yavas 1992; Wheaton 1990) The other group suggests that the longer TOM gets, the lower the price will be (see e.g. Taylor 1999).

As Sirmans et al. (2010) discuss, the relationship between price and TOM can be complicated. As buyer wants to minimize their purchasing price, the seller wants to obtain the highest possible price. The seller often knows the price range the apartment is reasonably priced. If a seller decides to choose a price from the top of this limit, the apartment will likely

have longer TOM because the number of potential buyers might decrease. The seller who selected the high price, may want to concede on the price, thus the relationship between price and TOM can become negative. However, the relationship may become positive. As Yavas (1992) argues, this can happen when a seller devotes more time for finding a potential buyer. By doing this, the seller can find a buyer who is eager to pay a higher price as a result of the seller's effort. However, the seller should consider whether they want to speed up the sales or wait for the potential buyer at the cost of waiting (Anglin et al. 2003).

On contrast, Taylor (1999) emphasizes that the longer a house is on the market, the more it raises doubts for the buyer that there exists a problem or a defect with the particular house. Taylor suggests that when a house has been on the market relatively long time, the potential buyer might become suspicious about the fact why the apartment hasn't already been sold. There might be several reasons for this, for instance, that the apartment is priced too high or that the previous potential buyers have detected a defect in a house.

Numerous studies report that overpricing and underpricing strategies prevent reaching the optimal TOM. Such as Springer (1996) proposes that overpricing increases TOM. Whereas Asabere, Huffman & Mehdian (1993) studied the relationship between price and optimal TOM. These authors created a linear programming model where listing prices are a function of the distribution of similar market listing prices. They stated that overpricing negatively affects the number of bids thus, the optimal time on the market won't be reached. Likewise, Li (2004) suggests that overpricing leads to longer TOM but does not pay off. Furthermore, Kang and Gardner (1989) show that overpricing is not a profitable pricing strategy. They used linear and log-log models to test house sales data from the US during the years 1982 – 1986. They found a positive relationship between overpricing and TOM. In addition, the findings indicate that the log-log performed better among these two functional forms. Their hypothesis aligns with Miller (1978), who argues that selling price increases when TOM increases.

When selling a house, the seller faces a simultaneous optimization problem because the seller wants to minimize TOM and maximize the sale price. However, this is not so straightforward in the primary residential real estate markets because the apartments may be launched for sale before the building starts, thus, there will be plenty of time to sell the apartments before their completion. This does not mean that the prolonged selling times are desirable, but the final sale prices are more valuable, and the developer's goal is not to

minimize the TOM. Actually, if the TOM is very short, it may indicate of a pricing error, that the apartment is put on sale with a too low listing price. Many researchers have investigated this relationship, and it seems that the list price plays a significant role, because it affects buyers' motivation and decisions. However, most of the studies do not focus on the primary residential real estate markets. They either consider both markets as a whole or only the secondary markets. This study will offer new insights by focusing only on sale prices in the primary markets.

2.5 Factors affecting Time-on-the-market

Asabere and Huffman (1992) studied the macro-determinants of TOM using house sales from Pennsylvania, United States, over a four-year period. They used both house characteristics and macro-variables to determine hedonic pricing model with TOM as the dependent variable. According to these authors, the number of bedrooms, bathrooms, and area does not explain the TOM in most of the constructed models. On contrast, interest rate, unemployment rate, and income, significantly explain TOM. The authors suggest that one-point increase in mortgage rate increases TOM by more than 42 %. The effect of mortgage increase is more powerful for larger four-room apartments, which increase by 119 % when the mortgage rate increases by one percent.

Some empirical works identify the factors affecting TOM by using survival analysis and hazard models (see e.g. Zhou et al. 2017; McGreal, Adair, Brown & Webb 2009). Zhou et al. (2017) take the hazard model approach to investigate the TOM of condominiums launched to presale in China. The study focuses only on newly built apartments which is unlike in many other studies. The study discusses that differences among apartments can affect TOM resulting that more preferred apartments are sold faster. Meanwhile, the TOM in homogenous projects should be more steady. The results propose that longer tom can yield to higher selling prices. Furthermore, the findings indicate that market conditions impact how the units are sold.

Kang and Gardner (1989) found that the size of an apartment does not determine the TOM. They also examine whether the mortgage contract rates affect to selling times of houses. They suggest that at times when interest rates are low, quick sales are more advantageous,

but when the interest rates are high, it's reasonable to wait and increase the selling time to gain higher prices. Mok (2003) argues that at times when prices are declining, TOM increases because sellers might hold "sticky" reservation prices. Whereas Jud, Winkler and Kissling (1995) suggests that price fluctuation in the market leads to a situation where a seller becomes hopeful about the movements in the prices and wants to wait to gain higher selling price which makes the selling time longer.

Several studies have shown that national and local economic conditions have an impact on TOM (see: Yang and Yavas 1995; Haurin 1988; Kalra & Chan 1994). Miller (1978) shows that TOM is a function of its attractiveness given the housing supply and demand conditions. Anglin et al. (2003) also support the fact that the market conditions can be discovered from the TOM model. When the market sees an increase in listings to sales, the TOM will increase for the sellers.

Li (2004) found that competitor actions significantly affect TOM through their effect on listing prices. The author also points out that the mortgage rate is a major variable influencing house selling time. Furthermore, other authors have constantly found a negative relationship between TOM and mortgages in cases when TOM is an indicator of the seller's searching cost. This is mainly because the mortgage rate is seen as a search cost for the seller. When the seller is a developer, the cost of search is the lending payment paid back. (Springer 1996; Li 2004) On contrast, mortgage rate can be seen as a cost for the buyer, and the lower the mortgage rates are, the lower the costs for the buyer. This in turn, increases the demand and reduces the TOM, and makes the relationship between mortgage and TOM positive. (Yavas & Yang 1995b)

Sirmans et al. (2010) study use meta-regression to study the relationship between the selling price and TOM. The author examines whether the coefficients of TOM are sensitive to income, sale of the year, model specification and location. The results show that the model is sensitive for all of these, except the location. Hui & Kung Yu (2012) found a relationship between changes in unemployment and price fluctuation and TOM.

Ong and Koh (2000) argue that apartment floor level affects TOM. They utilized the 2SLS method and found that apartments in lower floors are cheaper and take longer to sell, but also apartments in high floors have longer TOM. Furthermore, TOM is affected by price fluctuations in private and public housing markets. In contrast, according to Li (2004), there

is shorter TOM in apartments located on the higher floors. However, the author does not find explanatory power for house characteristics such as the number of bathrooms and bedrooms or view to pool to TOM. Also, Belkin, Hempel and McLeavey (1976) suggest that house characteristics does not explain TOM well.

Yavas and Yang (1995a) found that TOM is closely related on season-related variables, for example, they present that winter is a bad time to list high-prices houses on the market. They also found that volume in housing market activity affects TOM. The authors measure market activity by the number of listings at the local market at the time a house was listed.

In the light of previous research, there exist numerous factors that can affect to TOM. However, the results are contradictory, and the authors have gained different results using different data sets and models. This means, there previous studies do not provide clear support for the selection of variables for the models. However, it seems that besides the house price, there is evidence that market conditions affect TOM. Especially the relationship between the mortgage interest rates and TOM has been emphasized in many studies. This thesis will use house characteristics as well as market factors, such as interest rates as explanatory variables in the models.

2.6 Factors affecting house price

The previous literature emphasizes many factors that can affect house prices. One of the most important things affecting to house price is the house's structural attributes. These are the house characteristics, such as the age, number of rooms, and size. Numerous studies suggest that the area of the house is one of the most important factors impacting on the sale price of a house, followed by the number of rooms. (Garrod 1002) Multiple studies have also emphasized the effect of additional space, such as storage on the house price. (Forrest 1996)

Yu (2010) studied house prices using panel data from China’s 35 main cities. The results show that land supply and the area of vacant housing negatively affect prices, but mortgages have positive relationship with price. Rahmatian and Cockerill (2004) focus on studying how airport noise affects house value using hedonic models. They examined three functional forms which are linear, log-linear, and log-log. The fit of the linear model is not the best, and the R-square is significantly lower than in other functional forms. However, the linear

model coefficients are similar as in the log-linear and log-log models. The authors suggest that a non-linear model performs better than a linear form.

Multiple studies emphasize that the economic fundamentals have good explanatory power when explaining house prices (Yu 2010). Hwang and Quigley (2006) results show that the number of vacancies affects house prices and supply. In addition, they show that housing markets are local, and prices tend to increase as a consequence of the increase in employment rate and incomes. Capozza, Hendershott, Mack, and Mayer (2002) showed that real income growth, city size, population, and construction costs affect house prices. Also, Hort (1998) got similar results. However, some researchers have obtained opposite results. Gallin (2006) and also Mikhed and Zemčík (2007) results suggest that there is no stable relationship between house prices and economic fundamentals.

The available research from Slovakia is limited, and the studies of prices in Bratislava are found less. Rehák and Kácer (2018) studied the spatial structure of real estate prices in Bratislava. There is no official data on house prices in Bratislava, thus the authors used bid prices from house advertisements. The results indicate that Euclidean distance is the best proxy for distance to the city centre in hedonic pricing models. The study provide some additional findings as well. Larger apartments measured by area or rooms are more expensive. Extra features, such as furniture, increases the prices slightly and other additional services such as pre-paid parking showed a significant increase in prices. The height of a building did not seem to have such significant impact on the prices.

Dluhoš (2017) research results show that there is a significant negative relationship between distance from the municipality and closest large city and the municipality house price. In addition, there is a negative relationship between the house price and unemployment in a specific municipality. The author also points out that the house prices in Bratislava region are higher than in other regions in Slovakia.

Anglin et al. (2003) study results of the list price model show that list price is not varying along the season. Physical and locational variables explain most the list price. Additionally, the list price increases when sale volume rises. From this, it can be deduced that when uncertainty of increased market activity occurs, the seller increases their prices. The model also suggests that the seller do not decrease their sale prices when a level of competition and housing supply rises. Moreover, the increase in financing costs does not significantly impact

the list price, whereas vacancy has a negative and significant impact on the list price. In addition, in a situation when buyer and seller agree on a selling price, the discount of list price is uncorrelated with list price deviation.

Thanasi (2015) researched 1421 observations in Tirana, the capital city of Albania. The study adopts a multiple regression model to test hypotheses about the relationship between variables. Results show that parking accessibility, number of rooms, apartment square meters, view, and such characteristics affect apartment price. However, the location have the most significant impact on the value of an apartment. Wong, Chau, Yau and Cheung 2011 studied the effect of floor level and building height on apartment prices. They found out that floor-level matters.

It is clear that multiple factors affect the sale price of a house. The effect of supply on the prices seems to be universal. Most of the studies emphasize the impact of economic fundamentals on price, but some authors have found contradictory results, that there is no stable relationship between house price and economic factors. This study contains competitive intensity, competitor price, and interest rates to the model besides the house characteristics, which are the apartment size, floor, size of a balcony and terrace. Even though the literature highlights the importance of location, this study does not use it because all of the apartments share the same location.

2.7 Previous literature on competition in the residential real estate market

Papers that deal with residential construction projects, their sales, and competition are found much less. Hui, Liang, Wang & Wang (2016) takes a look at presale pricing in the residential real estate market under the game theoretic framework. The authors point out that pricing is seen as a complex decision-making process. The decision-making process should consider not only the traditional factors but also the nearby area. The authors' arguments that when considering the price for the project, the competitive environment should be taken into account.

Some previous studies use the real options framework to examine the project's competition. Bulan, Mayer and Somerville (2009) emphasize the role of competitors in construction processes. Their study focuses on how uncertainty affects the delay of investment and how

competition affects this relationship. This relationship is modeled using the real options method as competition affects the value of "option to wait". The results show a negative relationship between project development and idiosyncratic risk. This relationship is influenced by the competitive intensity which is described as a number of competing projects nearby.

Hui et al. (2016) examined the influence of developers' status and competitive intensity on presale pricing for residential real estate projects. They found out that the closer the competing projects are, the larger the price discount should be. In particular, these discounts to prices are 0.03 % in 3 km radius and 0.55 % in 1 km radius. Competing development projects nearby determine the competitive environment. The competition can be measured using competitive intensity. The location can be defined as the most important factor since it allows to observe the actual competing projects nearby.

Wang et al. (2016) study of the uncertainty, competition, and timing of land development indicates that the company's success can be affected by the competitor's decisions. The authors found that competition negatively affects investment timing thus, an increase in competing projects decreases the likelihood of development. In addition, their model found that competition affects the uncertainty of investment timing.

One famous research by Williams (1993) examines the competition in housing markets. The study applies games theory to real options. The results suggest that an increase in the number of developers will reduce the option value and shorten the waiting time before construction. The results are similar to what Grenadier (2002) observed. The author believes the investment is divisible, and the competition dramatically impacts exercise strategies. The study constructs a Cournot-Nash equilibrium framework that suits the best at the industry-level. The model shows that intense competition speeds up the timing of investments. In turn, Chu and Sing (2007) show that companies with significant competitive advantage don't need to accelerate the actions to occupy the markets.

Ooi & Thao 2012 examine a housing price adjustment when housing stock changes using vector autoregressive model. Their study focuses on the private apartment market in Singapore. The results indicates from "contagion" effect and suggest that when a new supply increases, the prices responds positively in the next quarter. The price adjustment is noted when the new houses are marketed, not when the construction is completed. The

market sees developers' choices to increase the supply as positive news of the future. However, there is a reverse, but insignificant effect in the primary markets. When supply rises, the new house prices tend to decrease. This shows that the rise in housing supply has different effect on prices of new and existing homes. According to the authors, there is herding action among homebuilders who tend to mimic other builders' timing when launching new housing units for presale. This can be seen when a supply change of quarter affects to the number of units released in the second quarter.

House price adjustments to supply changes depend on the housing supply's elasticity. If the elasticity is high, there is a greater effect of increased supply in price changes. If a perfect elasticity exists, a demand shock will be cleared out fast because the prices respond. In the housing market, the supply is not perfectly elastic. (Ooi & Thao 2012) Construction fluctuates with changes in demand, but the response time is slow because of the long construction time (DiPasquale and Wheaton 1994).

As a summary of this chapter, previous literature emphasizes that the developers should take into account the competition when making decisions. A number of papers have found a relationship between competition and real estate projects. Researchers have proved that competition can affect project sale prices, TOM, timing, and investment decisions (see e.g. (see e.g. Ooi & Thao 2012, Wang et al. 2016). These results highlighted in this chapter emphasize the importance of studying the competition effects on residential projects. Hui et al. (2016) have found that the closer the competitive projects are the more it has an impact on the project sales. Considering this, this thesis will take into account only the competition in the nearby area of the project.

Based on the literature presented in this chapter it is assumed in this thesis that as competition measured by the number of apartments increases, the final sale prices fall in the next quarter. This assumption is based on the competitive theory. Such as Ooi & Thao (2011) suggest that as supply increases, the prices fall. However, the competition effects on sales can be ambiguous. This means that an increase in competition can lead to lower prices to remain competitive advantage, or to higher prices due to increased market demand. Another assumption made in this study is that if an apartment is launched on the market when competition intensity is high, the TOM gets longer. The hypothesis formulation for time-on-the-market model is not straightforward because the previous literature has found contradictory effects of different variables on TOM.

3 Theoretical framework and methodology

This chapter presents the theory of hedonic pricing models. Rosen's (1974) study is used as a theoretical framework for hedonic models in this thesis. Chapter 3 is continued by the Porter (1980) five competitive forces framework, which is important to discuss in the context of what competition means for the developer. The process in which the seller and buyer are brought together is discussed in the last section.

3.1 Hedonic price functions theory

Hedonic price theory concerns differentiated products. Initially, it was developed for valuation of quality changes and improvement of qualitative indications (see e.g. Lancaster 1966). Hedonic pricing methods have its roots in 1939 when Andrew Court studied automotive industry. Court studied different factors and their impact on car prices. (Goodman 1998) Despite the fact that the theory was not originally developed for housing market purposes, it has become a popular method for analyzing housing markets (Laakso 1997, 24). Lancaster (1966) was one of the first who modeled goods as a collection of characteristics that form the heterogeneous product. In this model, consumers are searching for feature combinations that maximize utilities.

Theory of utility formation was developed further by Rosen (1974). Rosen presented the hedonic model where the property value is determined by the sum on property characteristics. This published article has been rewarding among housing market research (Laakso 1997, 25). Rosen (1974) suggests that markets construct of many products, each with different characteristics. The prices can be compared when these characteristics are broken down. Price differences can equalize various characteristics between products. This chapter is based on this Rosen article.

Hedonic price function is the core of hedonic price theory. A differentiated good is represented by a vector of its characteristics, $z = (z_1, z_2, \dots, z_n)$. In this case a house, these characteristics can be for example related to local environmental amenities or be structural, such as the number of bedrooms or bathrooms. Consumers can consume commodity that

includes vector z, and pay different prices for these commodities. The marginal price depends on the commodities consumed. The theory presents that the characteristics of a house can be measured. Every house has a market price P. Price of a house is a function of vector P(z) = P$(z_1, z_2, \dots, z_n)$, that combines the prices and characteristics. This price function is seen as a minimum price characteristic combinations, thus it describes the market price of a house. The main idea in hedonic pricing theory is to present how the price function P(z) is constructed.

3.1.1 Consumer's decision problem

Every consumer is supposed to purchase one unit of a differentiated product. It is also supposed that each z_i's are desirable characteristic and all consumers would want more of each feature. As a result, P$(z_1, z_2, \dots, z_n)$ is growing as the number the arguments increase. According to Rosen (1974), houses are heterogeneous and hedonic models are nonlinear. The model can be linear only in markets where no arbitrage opportunities exist. Consequently, nonlinearity is a reasonable assumption in housing market research. The consumer has utility function $U(x, z_1, z_2, \dots, z_n)$, where x represents consumption of other goods than housing and the price of x is set to 1. The utility function is assumed to be concave. The consumer decision problem can be presented as the maximization problem as follows:

$$\max U(x, z_1, z_2, \dots, z_n), \quad \text{s.t.} \quad y = x + P(z) \tag{1}$$

where y presents the consumer's income. The x and $(z_1, z_2, \dots, z_n)$ are observed as a result, satisfying both the budget constraint and first-order criteria that can be presented as follows:

$$\frac{\partial P}{\partial z_i} = P_i = \frac{U_{z_i}}{U}, \qquad i = 1, \dots, n \tag{2}$$

This means, each characteristic's marginal price is equal to its marginal utility. The buyer achieves the optimum by purchasing a product that has an optimum amount of each particular characteristic. The analysis is now continued by using a bid function of the consumer $G(z_{1,}z_{2,\dots,}z_n \; ; u, y)$, that includes the utility function and can be presented with the following equation:

$$U(y - G, z_{1,}z_2, \dots, z_n) = u \tag{3}$$

Essentially, the bid function presents a valuation function of house characteristics. $G(z; u, y)$ reflects the spending the consumer is ready to pay for values of $(z_1, z_2, \dots, z_n)$, when considering the given utility and income level. According to Greenstone (2017), the bid function is heterogeneous due to differences in the preferences of individuals. The following properties can be derived for the bid function:

$$G_{z_i} = \frac{U_{z_i}}{U_x} > 0 \tag{4}$$

$$G_u = \frac{-1}{U_x} < 0 \tag{5}$$

$$G_y = 1 \tag{6}$$

$$Gz_iz_i = \frac{U_x^2 U_{z_iz_i} - 2U_x U_{z_i} U_{zx_i} + U_{z_i}^2 U_{xx}}{U_x^3} < 0 \tag{7}$$

The bid function is increasing with respect to *z* while the marginal utility of *z* decreases. The bid function G stands for the amount the household is willing to pay for *z*. Other interpretation is that Gz_i can also be the reservation price of the consumer for an extra unit of a characteristic of z_i. The P(z) is the minimum price for the house. The utility is maximized when:

$$G(z^*; u^*, y) = P(z^*), \text{ and} \tag{8}$$

$$G_{z_i}(z^*; u^*, y) = P_i(z^*), \quad i = 1, \dots n \tag{9}$$

Where z^* and u^* stand for optimal values for *z* and *u*. The optimum is found when the P(z) and $G(z; u^*, y)$ are tangent to each other. This point is also a point where marginal price and marginal value cross each other. Figure 9 introduces decision-making position of two consumers with respect to characteristic z_1. The bid functions differ for consumers because they have different preferences and income levels, as already mentioned. Larger incomes lead to higher amount of each characteristic at the optimum point, if the P(z) is convex. Nonetheless, Rosen states that it is realistic to assume, that at the optimum, certain components are increasing and others declining in terms of income.

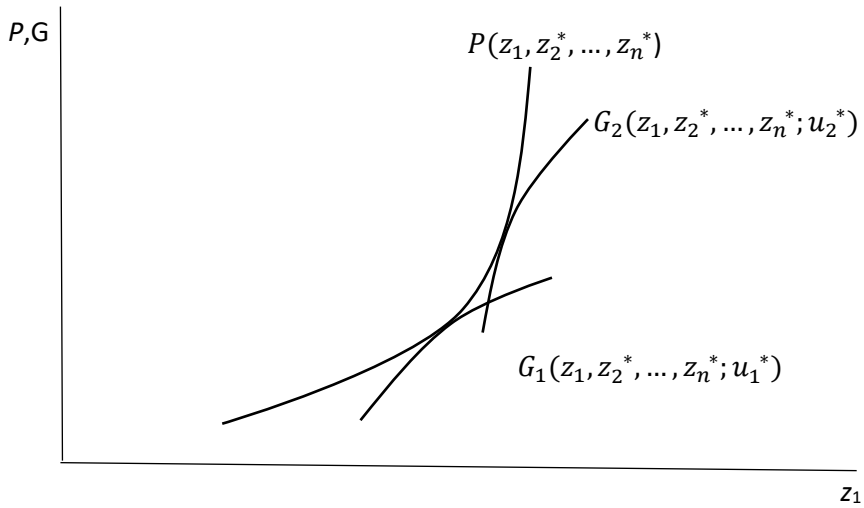


Figure 9: The bid functions of two consumers with respect to characteristics z_1 (Rosen 1974, 39)

3.1.2 Producer's decision problem

The companies can deal with production decisions in a similar way. M(z) describes the number of z-type products. It is assumed that joint manufacturing is impossible and companies production plant produces only one type of product. The total cost function for the production plant is denoted $C(M, z; \beta)$, where β stands for plant differences such as input prices or technology. The C is assumed to be convex and $C_M > 0$ and $C_{zi} > 0$. Plants maximize profit:

$$\pi = M\ P(z) - C(M, z_1, z_2, \dots, z_n) \tag{10}$$

It is important to note that the function P(z) presents unit price of the product of type z. It is assumed that markets are competitive, that makes the function P(z) independent of M. The optimal *M* and *z* can be presented as follows:

$$P_i(z) = \frac{C_{z_i}(M, z_1, z_2 \dots, z_n)}{M}, \qquad i = 1, \dots, n \text{ and} \tag{11}$$

$$P(z) = C_M(M, z_1, z_2, \dots, z_n) \tag{12}$$

The marginal revenue of characteristics *i* is the same as the marginal cost at the optimum. This means, the product is produced until price meets the marginal cost. The offer function $g(z_1, z_2 \dots, z_n; \pi, \beta)$can be determined for the producers. When the production volume of

each model is optimally determined, it reflects the unit price of a product that gives the company a consistent profit. The function can be solved by using the following equations:

$$\pi = Mg - C(M, z_1, z_2 \dots, z_n) \qquad \text{and} \tag{13}$$

$$C_M(M, z_1, z_2 \dots, z_n) = g \tag{14}$$

The offer function can be observed when removing the M and solving g as a function of z, π and β. It can be proved that it holds for g: $g_{z_i} = \frac{C_{z_i}}{M} > 0$ and $g_\pi = \frac{1}{M} > 0$. The optimum can be achieved when:

$$P(z^*) = g(z_1^*, z_2^* \dots, z_n^*; \pi^*, \beta) \tag{15}$$

$$P_i(z^*) = g_{z_i}(z_1^*, z_2^* \dots, z_n^*; \pi^*, \beta), \quad \text{i=1,…,n} \tag{16}$$

Figure 10 depicts two producer's offer functions in relation to one feature. The company offer function and hedonic price function are tangent to each other in the optimum. The marginal value and marginal price of characteristics i also cross each other that is also shown in the Figure 10. Production plants vary from one another in terms of parameter, indicating they specialize in manufacture of certain models.

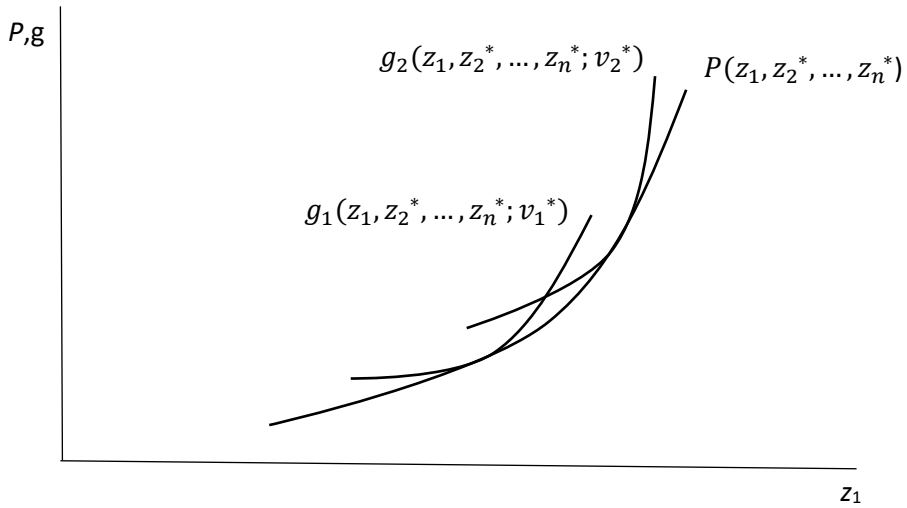


Figure 10: Offer functions of two producers with respect to characteristics z_1 (Rosen, 1974, 43)

3.1.3 Equilibrium

The consumer's bid function and producer's offer function are tangent in equilibrium and their common gradient at the tangent point is the same as the hedonic pricing functions gradient. Consequently, the function p(z) is the envelope of set of consumer bid functions and set of producers offer functions taken together, as presented in Figure 11. The equilibrium is established as a result of all customers' and producer's decisions. It is important to remember that in competitive marketplaces, each individual consumer and producer must accept the market price as it is.

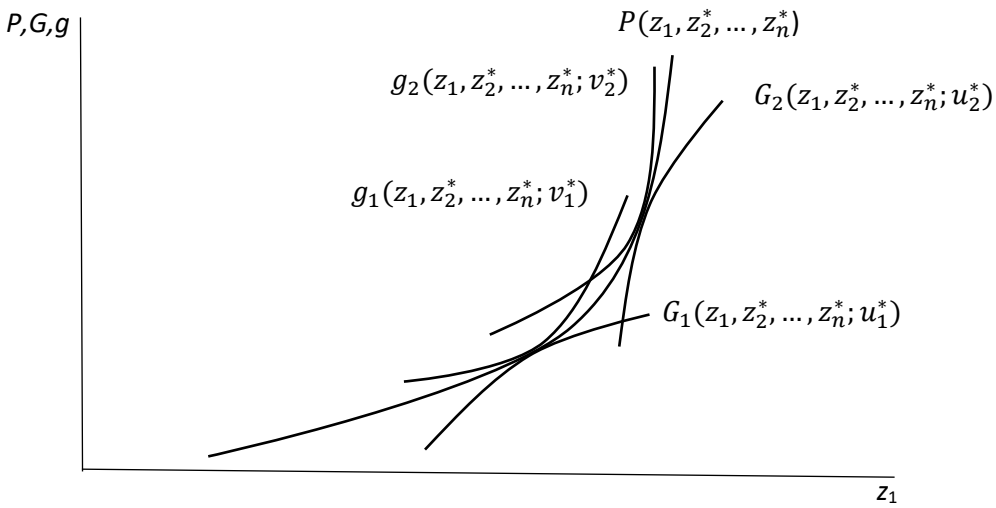


Figure 11: Equilibrium with two producers and consumers in case of one characteristics z_1 (Laakso 1997, 31)

3.1.4 Hedonic pricing model and functional form selection

Regression or hedonic models are widely accepted and a great number of previous studies use the method, which highlights its potential as an explanatory technique (see e.g. Can 1992). The method measures the relationship between house characteristics and its value, but similar models have also been used to analyze TOM (See e.g., Asabere & Huffman 1992). When estimating TOM, house selling time is a function of its different features, that can affect to this time (Sirmans et al. 2010).

The hedonic price model appears to be simple because it requires only specific data, such as the price of a house and a collection of housing features (Chau & Chin, 2003). Rosen (1974,

37, 47) presents some special cases where a price function P(Z) can be determined by the theory. However, these special cases require strong assumptions. Thus, there is no solid theoretical basis when selecting the correct functional form of hedonic regression.

Hedonic regression outcomes are dependent on the choice of functional form. According to Rosen (1947, 50), all of the characteristics of a house should be used as explanatory variables and the most suitable functional form should be selected. Empirical studies presents typical functional forms for hedonic price functions which are linear, log-linear and log-log (Bartik & Smith 1987; Halvorsen & Pollakowski 1981). All these three abovementioned forms can be estimated using Ordinary least squares (OLS). The OLS fits the model and determines the parameters using minimization on the sum of squared errors (Studenmund 2010). The linear model can be presented in the following form:

$$P = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \cdots + \beta_n z_n + \varepsilon \qquad (17)$$

where β_0 is the regression constant and $\beta_1 \ldots \beta_n$ are the coefficients and z_i describes the characteristics of an apartment i. The final term ε stands for the error term. (Laakso 1997) Follain and Malpezzi (1980) among others compared the linear and log-linear functional forms and found that log-linear have advantages compared to linear form. The log-linear regression model presents the response variable y on a logarithmic scale. The log-linear form can be presented as follows:

$$\log(P) = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \cdots + \beta_n z_n + \varepsilon \qquad (18)$$

Where log(P) is the natural log of price of an apartment. The log-linear model allows variation in the value of an apartment depending in part on the characteristics of that specific house. The coefficients have a simple interpretation because they can be presented as percentage change in the price of a house change in an independent variable. To add, log-linear form usually deducts heteroskedasticity, which is a commonly known statistical problem. (Malpezzi 2002, 20-21) The log-log model is a transformation of log-linear model, where the logarithm is also taken from explanatory variables. As Laakso (1997) states, the log-log model allows marginal price to either increase, decrease or to remain constant. Log-log- model can be presented as follows:

$$\log(P) = \beta_0 + \beta_1 \log(z_1) + \beta_2 \log(z_2) + \cdots + \beta_n \log(z_n) + \varepsilon \qquad (19)$$

When selecting the functional form for the model, the linear form is considered the worst alternative among linear, log-linear, and log-log functional forms, because the price function should be nonlinear in the case of nondivisible goods. However, the previous research does not create restrictions for the selection of the functional form. Many empirical papers assume a valid functional form without presenting any precise justification for the choice. (Laakso 1997, 41)

3.1.5 Challenges in utilizing hedonic models

The hedonic models have also received some criticism. Such as Chin and Chau (2003) have questioned the equilibrium assumption stated in the section 3.1.3. The real-world real estate market contains imperfections thus, the market equilibrium is implausible. The model's main assumption that consumers are aware of apartments for sale and hence adjust the consumption to new equilibrium based on their prices, preferences and income, is not common in real estate markets. Likewise, Halvorsen and Pollakowski (1979) arguments, the hedonic model can not be generally defined theoretically, as the hedonic price equation reflects the supply and demand effects. To add, the seller has more information of the apartment on sale than the buyer has, thus the information is incomplete.

The hedonic models assume homogeneity of housing. However, the houses differ from each other such as in terms of characteristics, neighborhood attributes, location, and other criteria such as building type (high-rise apartment, bungalow, condominium and such). Another underlying assumption is about perfect competition in the market, where there are many sellers and buyers. Because there are many purchasers looking for an apartment and many developers that provide housing, this assumption is justified. However, these parties can not significantly affect the prices as individuals. (Chau & Chin 2003)

One of the major issues in hedonic models concerns the choice of a functional form. Incorrect functional form can result in inconsistent estimates. Another issue that may arise in hedonic models is misspecification or omitted variable that is included in the model. (Chau & Chin 2003; Goodman 1998).

Considering the criticism hedonic models have received, there exist even more flexible functional forms such as Box-Cox transformation, that some of the researchers favors.

Laakso (1997) points out that the Box-Cox transformation has reached significant popularity in the literature because of its flexibility. The transformation allows to select between linear and log-linear models. For instance, Halvorsen and Pollakowski (1981) used this method to hedonic prices in order to eliminate the problem of using a form that is restrictive. The suggested functional form can be presented as follow:

$$P^{\theta} = \beta_0 + \sum_i \beta_i z_i^{\lambda} + \frac{1}{2} \sum_i \sum_j \gamma_{ij} z_i^{\lambda} z_j^{\lambda} \tag{20}$$

where P^{θ} is the price, z_i presents the attributes, $\gamma_{ij} = \gamma_{ji}$, P^{θ} and z_i^{λ} are the box-cox transformations, that can be determined using the following equation:

$$z^{\lambda} = g(z, \lambda) = \frac{z^{\lambda} - 1}{\lambda} \tag{21}$$

where the lambda varies between -5 and 5 (Soczewica 2021). Parameters θ and λ limits the form. When θ and λ are 1 and γ_{ij} are zero, the form becomes linear. When θ and λ are close to zero and γ_{ij} are zero, the form becomes logarithmic. Whereas when θ and λ are close to zero but some γ_{ij} are not zero, the form becomes translogarithmic model. (Halvorsen & Pollakowski 1989)

The box-cox transformation has also received criticism. One can argue that when estimating hedonic price models, reliable estimates should be developed for variables in price function, not just estimate models that explain the best the variation in prices. (Ohsfeldt 1988) Laakso (1997) argues that Box-Cox transformation does not significantly increase the explanatory power of the model itself. Box-Cox transformation can make the coefficients less reliable and more difficult to interpret hence the benefits are small compared to simpler functional forms (Greene 2008).

3.1.6 Model specification

The Grauss-Markov theorem famously states that the linear regression estimates are unbiased and have the least variance under a given set of assumptions. If a set of Gauss-Markov assumptions are met, the regression estimates are “Best Linear Unbiased Estimators” (BLUE). (Schmidt & Finan 2018, 147-148) The first Gauss-Markov assumption

that needs to hold that OLS is seen as BLUE is that the residuals follow the normal distribution. The Q-Q plot, which plots the quantiles for fitted values against standardized residuals, can be used to examine the normality in the regression. If the points fall on the line, the data is normally distributed. Normality assumption can also be checked using the histogram of residuals. The residuals should fall evenly below and above the mean in a bell shape. (Holm 2021) Additionally, the normality is tested using Shapiro-Wilk test. Shapiro-Wilk test is a statistical test that determines whether the data distribution differs from a comparable normal distribution. The results can be interpreted from the p-value and when p-value is larger than the 0.05, the distribution is not significantly different from a normal distribution. However, when test p-value is significant when p-value is below 0.05, the distribution is significantly different from a normal distribution. (UC 2022)

Linear regression assumes that there exists a straight-line relationship between response variable and predictors. If such relationship is not linear, the prediction accuracy can decrease and the conclusions from the fit are questionable. (Gareth, et al. 2013, 92) This linearity assumption is the second Gauss-Markov assumption that make OLS BLUE. As discussed earlier, the use of logarithm transformations of variables can help make the relationship more linear. Linearity assumption can be verified by plotting the predictor variable against fitted values. If there seems to be a pattern such as a strong curve, a more complex model than linear might be needed. (Casson & Lachlan 2014, 593)

The third Gauss-Markov assumption is that error terms have constant variance, $Var(\varepsilon_i) = \sigma^2$. Heteroskedasticity can be observed by examining the plot of residuals against fitted values. If there exists a funnel shape when plotting the residuals, heteroskedasticity might be present, which means, the residuals are not normally distributed and have non-constant variance. If the plotted residuals tend to increase with fitted values, one possible solution is to transform response variable Y to logarithmic. Heteroskedasticity can lead to forecasting error and also confusion in the interpretation of the model. Heteroskedasticity can be tested by using Breusch-Pagan test. (Gareth et al. 2013, 95; Singla& Priyanka 2019, 1040).

The presence of multicollinearity can cause problems in regression analysis because it may be hard to distinguish the individual effects of collinear variables on the response variable. This is also one of the Gauss-Markov assumptions to make OLS BLUE. If predictors variables are closely related to each other, collinearity can be observed. The presence of multicollinearity causes standard error to grow, decrease the accuracy of the estimates of

coefficients, or overestimate the R2. Possible multicollinearity can be observed by creating a correlation matrix or computing variance inflation factor (VIF) values. (Gareth et al. 2013, 100-101; Singla & Priyanka 2019, 1040) VIF measures the degree of multicollinearity in independent variables. If the VIF value exceeds the threshold, which is often 10, the multicollinearity problem is observed. The smallest value VIF can get is 1, which means the absence of collinearity. (O'Brien 2007, 674-683) If multicollinearity problems are ignored, the parameter estimates can be inconsistent and unreliable (Laakso 1997, 46).

One typical problem in the hedonic models is the large number of variables that can affect to house price. As the number of these variables increase, the possibility of multicollinearity problems may rise. Multicollinearity problem is a common issue in hedonic price models, for example, usually as the size increase, the room numbers of a house increase. To solve the problem of multicollinearity, one or more variables can be removed. It must be noticed that removing a variable from the model does not always make any improvement. Step-wise regression could be used to eliminate the presence of variables with high VIF values (Gareth et al. 2013, 101; Dormann et al. 2013).

Further, it is important to check whether data contains outliers. Outliers are data points that do not fit the model. Leverage points are outliers with respect to predictors. Outliers can be identified by plotting the residuals. (Gareth et al. 2013) According to Blatná (2006, 1), there can be two types of outliers. The good ones improve the precision of the regression coefficients, whereas the bad ones are located far from the regression line, and they may reduce the precision of regression estimates. Outliers not only cause implications for interpretation of the fit, but might also cause a decline in R-square and an increase in RSE. (Gareth et al. 2013, 97-98)

Cook's distance can be used to detect outliers. Usually, four divided by the number of observations is used as a cut-off value, and if a data point is larger than this, it is considered as an outlier (Bollen & Jackman 1990). Each model run in this thesis are examined using residuals versus leverage plot, from which one can see if the model contains high leverage points that might be influential. Since the data is small, each high leverage point won't be automatically removed, but is being tested whether the removal impacts on the regression results. As Gareth et al. (2013) suggests, an outlier may or may not have an effect on the least squares fit, thus removing it, it can be checked whether the model can be improved.

One method to evaluate that the model does not have omitted variables and is correctly specified is to use the Ramsey's RESET test. Ramsey (1969) introduces a general specification test to examine whether non-linear combinations of fitted values can be used to explain the dependent variable better. The null hypothesis is rejected if the model has incorrect functional form, or the error terms correlates with regressors. If the null hypothesis is not rejected, the powers of fitted values can't explain more the dependent variable (Volkova & Pankina 2013).

The regression model can be evaluated using R-squared metric, which is the most commonly used measure of goodness of fit. R-square measures the percentage of the variance in dependent variable that can be explained in the model. This means, that if the value is higher, the more the model can explain the variance in the dependent variable. The R-square can get values between 0 and 1.

Malpezzi (2002) suggests that use of dummy variables can be useful and bring more flexibility to the model. This is because if using for example, room numbers in a regression, it limits percentage change in value from one-room house increase to three-room house to be same as percentage change in value increase from one-room house to six-room house. Also, Ong & Kao (2000) suggest that dummy variable is preferred to allow non-linearity in the model.

3.1.7 Summary of hedonic models

The hedonic regression model is one of the most used tools to estimate housing markets. Despite the widespread usage of hedonic models, researchers are not agreed which functional form should be used. Economic theory offers little guidance on constraints on the functional form. However, it seems universal that log-linear form typically performs better than linear functional form. There are several key assumptions that need to hold so that OLS is seen as BLUE. In this thesis, different functional forms are compared in order to find the model that fits the best. As a result, this study takes advantage of linear, log-linear, and log-log models. The log-linear model presents the response variable y on a logarithmic scale. The study also utilizes a log-log form, where dependent variable, as well

as some of the explanatory variables are logged. Furthermore, this study takes advantage of dummy variables to bring more flexibility to the models.

3.2 Competition

Each competitor's decisions are affecting to other companies in a market. One can argue that competition speeds up the investment decisions and competitors compete over the timing of construction. (Wang, Tang & Jua 2016) However, companies should always consider the price in reference to the competitive environment. Clearly, the competitive intensity significantly impacts the supply curve in the area where the project is located. (Hui et al. 2016, 1204) External competition in the housing market occurs in markets where substitutes are sold. The apartments sold by different developers are substitutes for each other and make the owner-occupied housing. (Atterhög & Lind 2002)

Increased competition most likely leads to lower house prices (Kuryj-Wysocka, Radosław 2013). For example, in Sweden, the house prices are approximately 20 percent above the average prices in the EU. Swedish competition authority sees this as a lack of competition, especially in the construction sector and rental housing market. (Swedish Competition Authority 2002a; Swedish Competition Authority 2002b)

Porter (1980) offers an analytical framework for understanding the competitors and industries better. The framework shows how to define the industry attractiveness and its causes and it directs the firms to recognize the most critical aspects of industry structure to achieve long-run profitability. Porter introduced five forces that a firm should consider when acting against market competition. These five competition forces include a threat of new entrants or substitute products or services, bargaining power of buyers and suppliers, and competition between existing companies (Figure 12). The framework shows how competitive strategy can affect these five forces. Besides, it defines how one can analyze the competitors in the industry and impact to their behavior. (Porter 1980, 2-7)

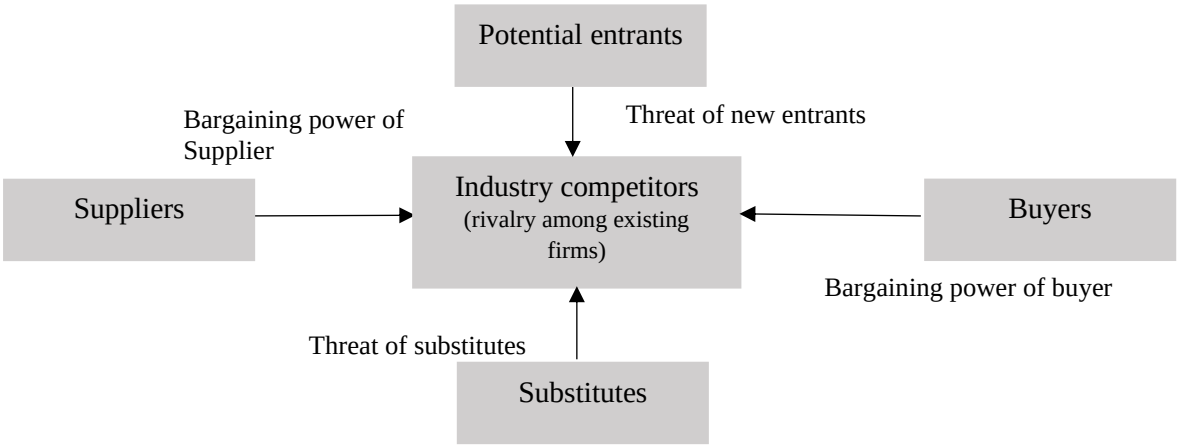


Figure 12: Five Competition Forces Approach (Porter 1980)

The forces determine the profitability of an industry because they affect the return of an investment. The prices are influenced by the threat of substitutes and the buyer's bargaining power. Buyers can also have the power to require more costly services which will increase the costs. Slovakia currently faces a lack of apartments while demand is high (Liptáková 2020), which makes the seller's markets. Consequently, buyers have worse bargaining power. However, the buyers are individuals thus do not negotiate together the best deals. The buyer can have bargaining power for example, in situations when the purchaser is buying a substantial part of apartments in sale. (Zhang 2010) In the case of new development, tendering is not as typical as in the secondary markets. In this particular project, the prices are list prices, and the buyers don't have bargaining power during sales negotiations.

Suppliers with bargaining power can affect to the costs through the raw materials (Porter 1985, 5). The supplies in the real estate industry are building materials, equipments and land. Zhang & Ren (2006) suggests that land provider has highest bargaining power, followed by special material producers. Suppliers of general materials do not have significant bargaining power.

The rivalry in the industry can affect both to prices and costs, depending on its intensity. The threat of entry can limit the prices and investments. (Porter 1985, 5) Different factors define the level of rivalry such as cost of exit, market concentration and maturity of the housing industry (Zhang 2010). Since the construction processes are lengthy and contain various phases, including the planning and other preparations (Healey 1994), exiting from the market is costly. All functions in construction industry require significant capital investments, thus

the rivalry in such market is smaller than in other markets. Moreover, the lower the concentration ratio, the more challenging the competition. (Zhang 2010)

The competition between rivals is the competition between similar houses in the same area. Competitive intensity in each local market depends on how many companies acts in the area. This force is seen as the most important. (Zhang & Ren 2006) In 2020, the construction sector totaled 123 905 companies in Slovakia (European commission 2021). It's clear that most of these companies are still not competing with the project analysed in this thesis. For instance, most of the companies are very small-scale enterprises. According to Bošácky (2017), there are approximately 10 major developer companies which are competing of the Slovak development market. These developers sold most of the housing units in Bratislava in 2017 (Suchý 2017).

As a summary, the Porter's theory suggests five competitive forces that define the competition intensity of a trade. The enterprises can face different natures of competition, in which the competition among rivals is seen as the most important competitive force among the five alternatives for the developers in residential real estate market. As the competition is fierce, most of the firms gains lower profits, *vice versa*. The buyer's bargaining power is not seen as such impactful threat since there is a lack of housing in Slovakia, that makes the seller's market. This study considers the competition within existing rivals, but does not take a stand whether new projects are provided by existing rivals or new market entrants.

3.3 Search theory

Search theory is often used to explain the process of the time required to match a seller of a house with a suitable buyer. Search theory explains the trade-off relationship between the list price and time a house takes to sell. The seller wants to maximize the discounted present value of realized profits from sale by selecting a price that strikes the balance between marginal costs of searching and marginal advantages of accepting an offer. On contrast, the buyer has a similar problem when looking for a home that maximize the utility given by housing services available in particular house. (Knight 2002)

The seller have a house in sale at a certain price. The both parties can value the house differently. These valuations are private information for each party. (Yavas 1992, 535)

According to search theory, both the seller and buyer of a housing unit face search costs as a result of imperfect information. The seller searching cost can be such as maintenance expenses, selling expenses and additional mortgage payments. (Li 2004)

Knight (2002) discusses about the process of bringing a buyer and seller together. The seller competes with other price offers from which the buyer seeks the most attractive deal. As a results of this searching process, the buyers and sellers are matched in bilateral pairs. When these parties meet, the buyer reviews the house and makes a decision whether to buy a certain housing unit. Some research papers emphasizes the role of listing price in this searching context, such as Mayer (1995) propose that the choice of listing price has an effect on the arrival rate of buyers. The list price itself, is a function of the market price for the houses, given its characteristics and availability of housing.

The list price acts as a strategic element from which the buyer makes conclusions. A buyer pays attention to a potential house based on the listing price and possibly makes an offer. In a secondary markets, the seller can make a counter-offer which the buyer can either accept, reject or make a new offer and continue the negotiation process. This process can end up to closing the deal or the house stays on the market. (Dubé & Legros 2015)

As a summary, a search theory is frequently used to describe the selling process in real estate market. Trade-off between seller's list price and it's marketing time is explained by search theory, as the seller wants to maximize the price in the shortest amount of time. However, this is not straightforward in primary residential markets. The apartments are launched to sale at an early stage, meaning that the developer does not want to make the sales quickly at the cost of higher profit. However, it must be noted that there is a cost of waiting for both parties, thus an optimal waiting time to maximize profit should be found for the seller. Search theory emphasizes that the buyer and seller meet when their incentives meets, and if the apartment remains unsold, the seller might want to change the pricing strategy to attract more buyers. This thesis assumes the higher prices will lead to longer waiting time for the seller.

4 Data

The next part of the thesis focuses on presenting data and variables used in the study. The section starts with a descriptive analysis of the project after which its competitive environment is visually explored. After descriptive analysis, the data collection procedures and variables are introduced.

4.1 Descriptive analysis of the project

The final data includes nine construction phases of the selected project, where each phase is one multi-story building. All the apartments have been sold from the first four phases. Most of the apartments have been sold from phases five to eight. Phase nine apartments were launched for sale near the date the data was gathered, so not even half of its apartments are sold. This thesis focuses to examine only the sold apartments. Figure 13 describes the sales and new launches of apartments of the project over time. The line represents the average sale price of apartments in square meters in euros. The average prices have been growing steadily over the period. The most intense price growth was seen in 2021. The bars colored with light blue in represents the amount and time the project launched their apartments into sale.

The new supply tends to enter to markets in distinct batches. In 2016, 2019 and 2021 one project each year launched their apartments into market. On contrast, in 2017, 2018 and 2020 two phases launched their apartments into markets each year. The peaks in sales tend to be at times the project launches a phase on sale. However, this is not in all of the cases. Such as, after the fourth phase launch in June 2018, there wasn’t any peak right after the launch date. On contrast, in April in 2017 there was not any sale transactions even tough the apartments were launched in March. This is because the process from reservation until sale can take some time. Furthermore, as there are so many apartments launched on sale at the same time, the buyers might think they don’t need to hurry with the sale process and reservations because there are plenty of apartments available.

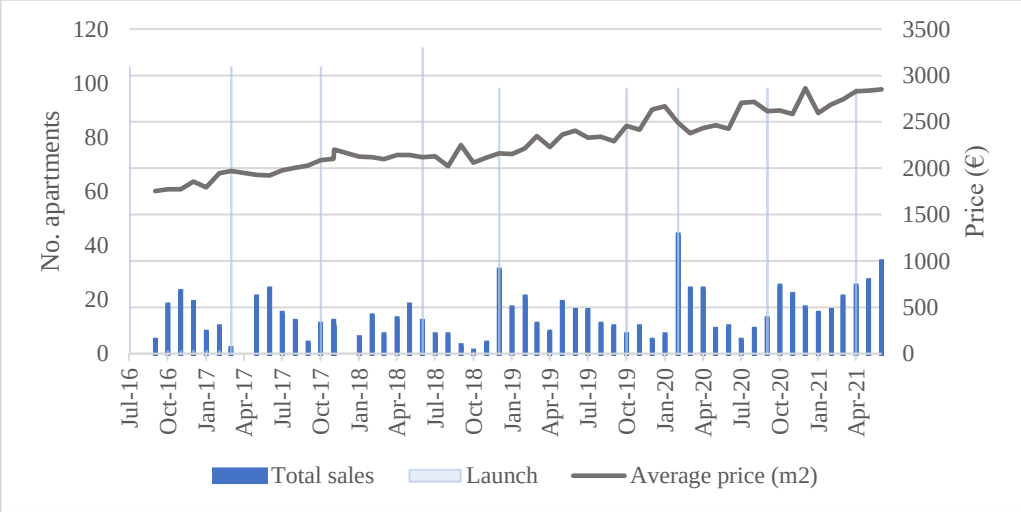


Figure 13: Price development of project apartments (price in square meters) over time and total sales

Figure 14 shows the distribution of sale price by phases on the left. The phases are in chronological order. The phase 9 have the highest average sale price followed by phase 8. However, the median price is highest in the phase 6. The prices have been rising over time and each new phase has higher selling price than the previous one had. There is a lot of variability in the prices in each phase, but that is mainly because of the different size apartments. Other factors can affect the price as well, as discussed earlier. The longest TOM have been in the phase 4 and the shortest in phases 3, 7, 8 and 9 (14 days), as shown in the Figure 14 on the right side. On average, the phase 5 have had the longest selling time. The latest phases have shorter selling times mainly because they have been launched for sale closer to the date the data was gathered and not all of the apartments are sold.

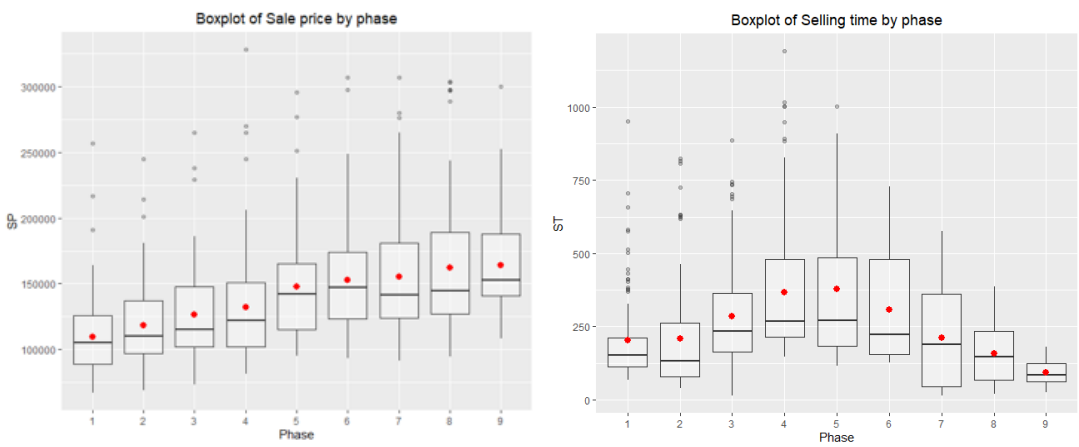


Figure 14: Boxplot of the sale price and TOM by phase

Histograms of TOM in each phase are presented in Appendix 1. The Appendix 1b illustrates the histograms in the same axis to make different phase performances easily visually comparable. The selling times differ from each other a lot. Since the last phases have been on sale relatively short time, their selling times can not be that long than in the first phases. The figures can reveal if some of the phases show strong initial performance. The second phase shows signs of outperformance since it has strong initial sales with a high frequency of sales after the sales started and only a low number of sales with long selling time. Strong performance in phase 2 can be for example, an effect of the marketing efforts during the phase 1. This would mean, the early-stage marketing efforts showed its full potential in the second phase, which can indicate of the spill-over effects.

Similar conclusions can be drawn from phases 1 and 7 since their histograms are right-skewed and a great amount of the apartments are sold quickly. On contrast, phases 5 and 6 have small peaks at the end of selling period, phase 5 has relatively many sales with more than 800 days and phase 6 more than 500 days. To conclude, in most of the cases, the trend in selling times seems to be declining over time. Otherwise, there does not seem to be other apparent patterns in selling times between the phases.

Selling prices are also presented by floors in the following boxplot in Figure 15 on the left. The highest average prices are in eight and twelfth floors. That is reasonable since in eight phases are eight storey-buildings and one phase is 12-storey-building. This indicates that top floors has the highest prices. On contrast, the lowest selling times on average seems to be in the bottom floors. Figure 15 right side presents how the selling times differ among the floor levels on the right side. The average selling time is growing along with the floor level. In the top floor there is a lot variance, some of the apartments are sold within a year, and the longest selling times takes almost 1000 days. The smallest variance is selling days is in 11th floor. The shortest selling times on average are on the third floor.

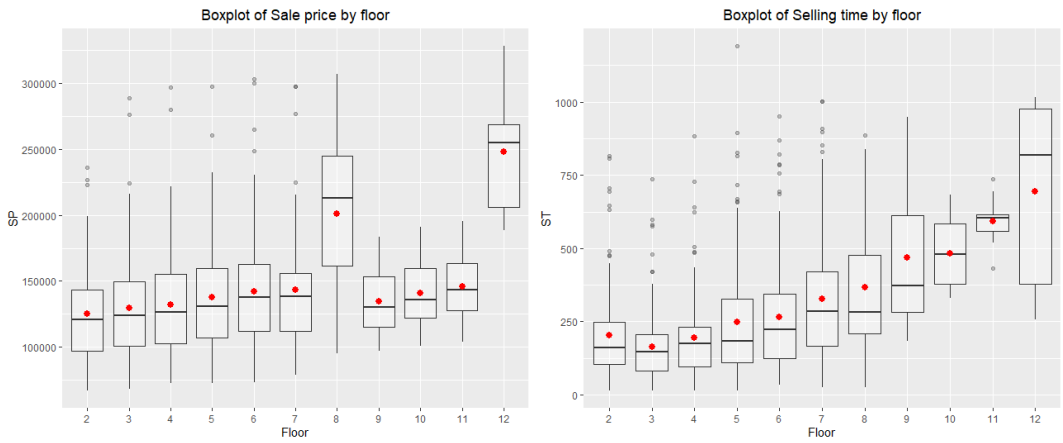


Figure 15: Boxplot of Sale price by floor and Selling time by floor (data contains 8- and 12-storey buildings)

There are also differences in selling times between the type of an apartment as introduced in Figure 16. On average the 2-room apartments have the lowest selling times, followed by 3-room apartments. The selling time for studios and 4-room apartments are almost the same, but the 4-room apartments has slightly longer selling time.

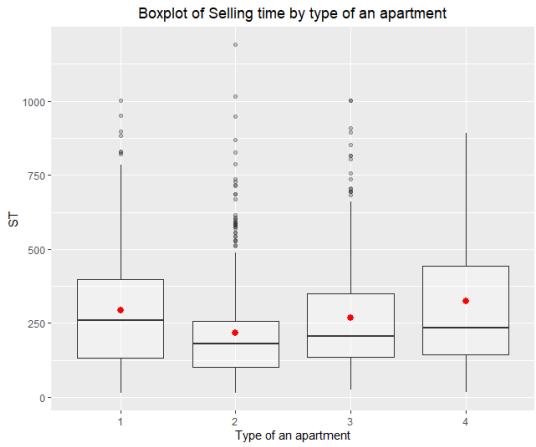


Figure 16: Boxplot by selling time by type of an apartment

4.1.1 Characteristics of competition

This chapter aims to introduce the competitive environment in BA II district in Bratislava. The information captured in this chapter figures are gathered from Lexxus research company that provides market analysis reports (Lexxus 2016-2021). Figure 18 presents the number of available newly-built apartments in residential projects at the end of each quarter in BA II during 2016 – 2021. On average, there have been approximately 707 available apartments in each quarter. The number of available apartments reached its top in the second quarter of

2017. On contrary, it reached its bottom in the third quarter of 2019. Average tax-free prices of available apartments in residential projects has been growing over the period. There has been a small abnormal peak in the second quarter of 2020.

Figure 17 describes the structure of available apartment units in residential projects in BA II area over time. Traditionally, the category of two-room apartments has maintained the leading position. The category of three-room apartments are ranked as the second largest group in each year. 1-room apartments are available much less, and the share of larger apartments is even smaller, between 0-2 % each quarter.

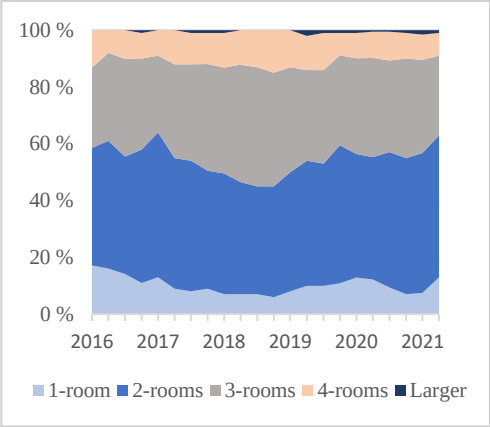


Figure 17: Share of available apartments by apartment type over time

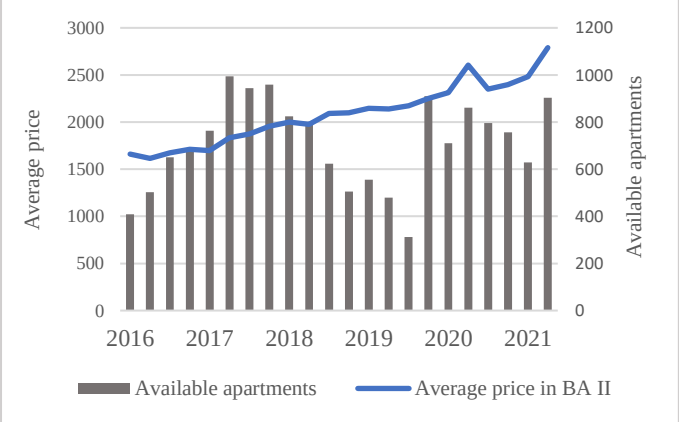


Figure 18: Available newly-built apartments and their average prices across all sizes in BA II

The number of new residential real estate projects or their phases that launch new apartments into market in Bratislava II area are presented in Figure 19. The project phases that are analysed in this thesis are also included in these numbers with the blue color. Projects that launch only 1 or 2 apartments were eliminated because those will distort the overall picture of the number of new projects. In three quarters there were eight new projects. On the contrast, there were no new projects launching new apartments on sale in the last quarter of 2018. In 2016 and 2017 the number of projects has been mostly growing, but after that the number of projects has been highly volatile.

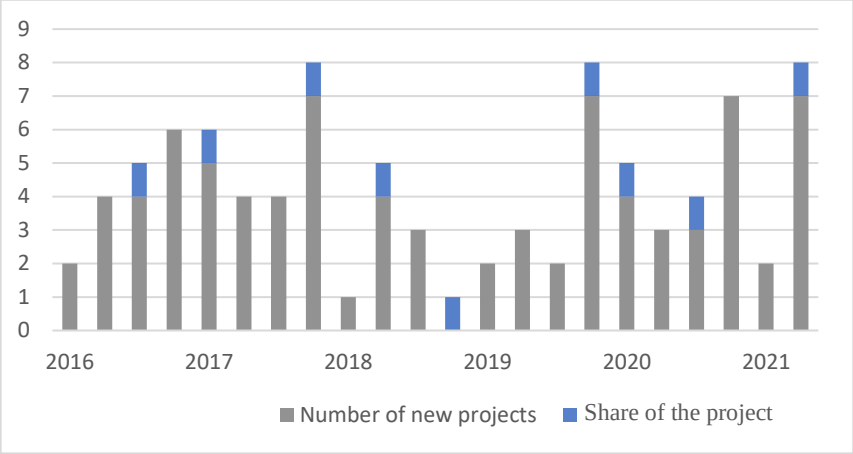


Figure 19: Number of new projects that launched apartments into market in BA II

Figure 20 presents the number of new apartments that the competitor projects have launched to markets. The number of new launches have been mainly growing in 2016 and 2017, after which the number have been decreasing. The number of apartments within the projects rose considerably in 2019 last quarter because there were 1310 new apartments offered for sale. This was mainly because of couple new large project with great number of new apartments. After this, the number of apartments launched to market has been rising. The dispersion in quantities launched differs a lot from each other. The last quarter of 2018 did not record any new apartments to sale. On average, 270 new apartments have been launched to markets by competitors in each quarter in BA II.

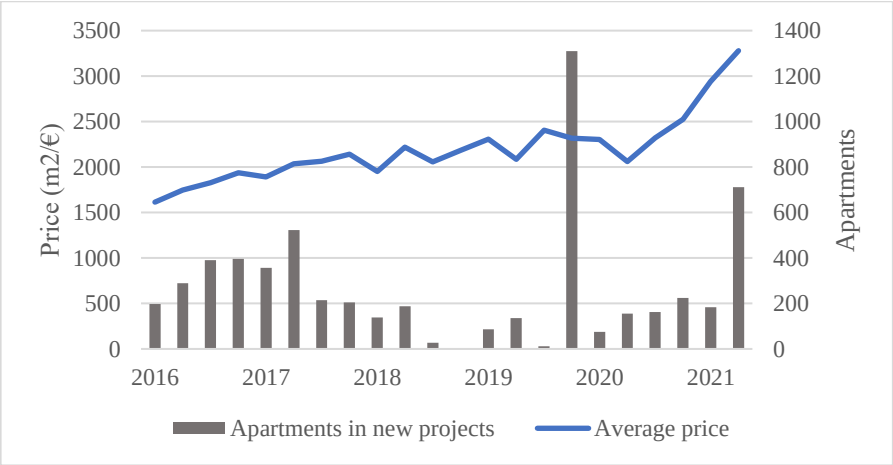


Figure 20:Number of apartments launched to markets and their prices in BA II

4.2 Deriving the dataset

The empirical dataset employed in this research is transaction sales data provided by one major construction company in Slovakia. The data set contains 936 apartment sales from one construction project from the third quarter of 2016 to the last quarter in 2021. The data contains both sold and unsold units. The final sale prices and selling times of unsold apartments are not known, thus they are not considered in this analysis. The observations of sales data include unique information of each sold apartment. This information includes variables such as listing price, sale price, sales status, date for marketing, date of sale, information of size of a balcony and terrace, room numbers, floor, size, total size, phase number, stairway, and building completion date. However, not all of these variables are used in this thesis.

To include the effect of competition in the nearby area, data of competitors was collected. There is no available information of competitors' each sale transaction in Bratislava residential real estate projects. Therefore, quarterly data of competitors' sales were used. Data of competing projects is gathered manually from residential real estate market reports made by Lexxus a.s. The company provides quarterly analysis from Bratislava residential real estate markets. Data is missing from third and fourth quarters of 2021, thus the observations that were sold in the last quarter of 2021 and those which are launched in the market in the third quarter are removed from the data to avoid missing values in the analysis. In consequence, the final data covers sale period from third quarter of 2016 to third quarter in 2021. The sales in the last third quarter can be included to the analysis despite of missing values in competition variables because the lagged values of them are used, as discussed later in this section.

Interest rates are also included in the analysis to characterize the market condition. All of the market factors are combined with the sales data, by creating new columns for each new explanatory variable. Each sale transaction is matched with market variables using newly created marketing quarter and sales quarter columns. The marketing quarter column is created using the marketing date the unit was launched to sale and the sales quarter using the actual selling date. Data collection procedures are presented in figure 21.

The variable selection is done based on the variable relevance in the study and in the light of previous research. Some of the irrelevant variables are removed. The variable removal is

done if assuming the variable won't bring much insight to analysis, such as a stairway or building completion date. After data processing, there are left 865 observations and 9 construction phases. One competitor variable contains two missing values, which are discussed later in this section. Otherwise, the sales data does not contain any missing values. As the study considers only newly built apartments, each apartment is sold only once.

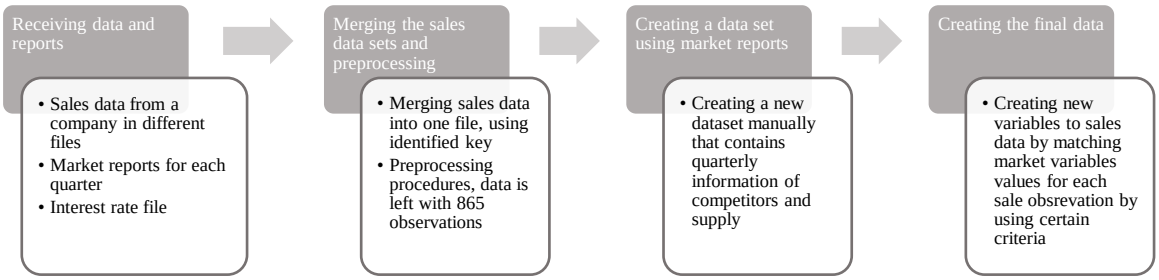


Figure 21: Data gathering process and final data formation

4.2.1 Variables

This section introduces the variables used in this study in more detail. Table 5 describes the mean, median, maximum, minimum, median, standard deviation, and skewness of variables. Most of the variable skewness is positive, which indicates the distribution tail to the right. If the value for skewness is negative, the distribution tail is to the left. The closer the normal distribution of the variable is, the closer to zero is the skewness value. (ESH 2022)

Apartments are launched in sale with a listing price that contains the price of a parking place and storage. The data set used in this thesis does not include information on whether the price includes the parking place or storage or not. This means the list prices are not feasible to use because the composition of list price is not fully known. This means, for example there might be apartments X and Y. The apartment X might have lower price because it does not have parking nor storage, whereas apartment Y has parking place which is why the price is higher than apartment X. However, the final sale price does not include the price of parking and storage, which is why the selling price is used as a target variable in this study. The final sale price and list price are usually the same, which is why it is acceptable to use final sale price in the study. The final prices are tax-free.

The histogram of price shows that the variable is significantly right skewed (Figure 22). In order to make the sale price more normally distributed, a logarithmic form of the sale price

is used. The transformation makes the variable closer to normal distribution. The average price for apartments is 136 950 euros, while median is 128 848 euros as shown in Table 5.

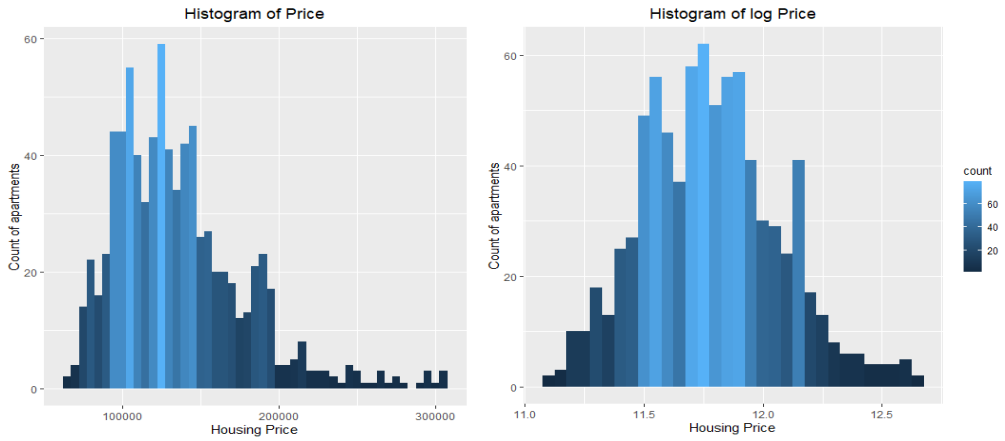


Figure 22: Histogram of sale price without transformation and after taking the logarithm

The area of a unit is measured in square meters and is captured in variable *Size*. Actual size of an apartment is used, and the total size is removed since it also contains the size of possible balcony or terrace, which are already covered in other variables discussed later in this section. It can be seen from Table 5 that the average area of apartments is found to be slightly below 60 m^2. To control for possible outliers in this variable a scatter plot is made. A plot of size against the prices indicates of a linear relationship and there seem not to be any heavy outliers, although one apartment is relatively large compared to others (Appendix 2).

Most of the apartments in the project are two-room apartments. The share of four-room apartments is very small, approximately 7 %. The final data contains 130 one-room apartments, 396 two-room apartments, 282 three-room apartments, and 57 four-room apartments, as shown in Appendix 3. This information is captured in *Rooms* variable. The average room type is 2.31, and the skewness for this variable is 0.15 which is close to zero. This indicates that the distribution is quite normally distributed. New dummy variables *Room1*, *Room2*, *Room3* and *Room4* for each type of apartment are created. If an observation belongs to a certain apartment type group, it will get value of 1, if it doesn't, the value is 0.

All the apartments have either balcony or terrace, or even both. In the final dataset, 21 apartments have both balconies and terraces, 34 have only a terrace, and 810 apartments have a balcony. The data contains dummy variables for balcony and terrace separately. The size of the balcony and terrace are captured in variables *Balcony_size* and *Terrace_size*.

Terrace and balcony dummy variables are highly negatively correlated (-0.78). Also, Terrace size and balcony are negatively correlated, and terrace size and terrace are positively correlated. Because of the high correlation with balcony and terrace dummies, they are not used further in this analysis; instead, their size is used. In addition, since each apartment has either balcony or terrace, the dummy variables won't bring much value to the analysis.

The size of a terrace and balcony is presented in square meters and if the apartment does not have balcony or terrace, the size is zero. The average balcony size is about 6 square meters, whereas average terrace size is only 2 square meters. This is because for most of the apartments terrace area is zero, which means they have only balcony and no terrace. This diminishes the average terrace size. For those apartments that have a balcony or terrace, the areas differ between 2,6 – 14,43 m^2 for the balconies and 4,08 – 85,98 m^2 for the terrace.

The floor in which the apartment is located is also known. Eight of the buildings have eight floors and one building has 12 floors. The distribution of apartments over floors is presented in Figure 23. It is reasonable to assume that there are price differences among the floors. Because of the different heights of the buildings, the comparison of, for example a unit in the floor eight in eight-story building and a 12-story building is not feasible. It is assumed that floor eight has higher demand and prices in the eight-storey building because it is the top floor, compared to same floor number in higher building.

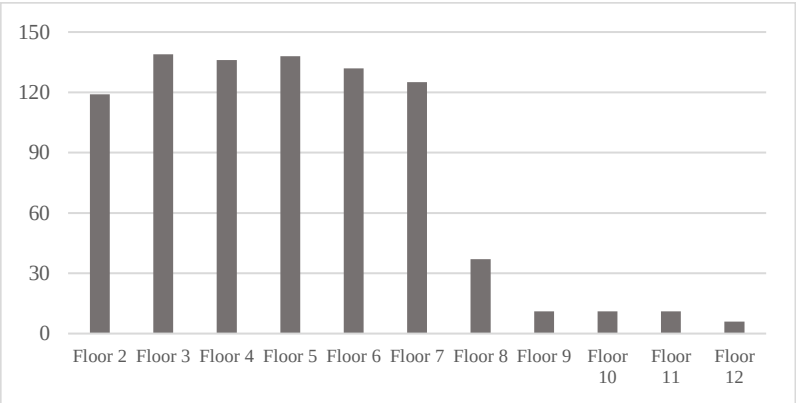


Figure 23: Distribution of apartments over floors

As suggested by Ong & Kao (2000), dummy variables are created for *Floor* variable to consider the different floor heights of the buildings in order to capture possible price and selling time differences on different floors. The floor levels are grouped into four separate groups, as presented in Table 3.

Table 3: Dummy variable creation for floors

Floors in total	Ground	Mid	High	Top
8 floor building	2, 3	4, 5, 6	7	8
12 floor building	2, 3, 4	5, 6, 7, 8	9, 10, 11	12

New dummy variables for seasons are created based on the launch month. The observation is assigned to dummy *Spring* if it was launched to market between February and April, *Summer* if between May and September and *Fall* if between October and January. This allows comparing different seasons, whether the launch schedule has an impact on the sales.

Selling time, also known as TOM, is captured in the variable *TOM* and is calculated by the difference between the date the apartment is listed on sale and the day it was sold. Selling time is defined as the number of days. The average selling time has been 253 days, whereas the median is 194. Four apartments are sold over 1000 days, which is a comparably long time. Histogram of selling time shows that it's significantly right-skewed (Figure 24). There are many observations with comparably short selling time, whereas there are several observations with extremely long selling times. Log transformation of the variable makes the distribution closer to normal distribution. TOM is expected to have a positive coefficient in the selling price model because when TOM gets longer, there is a higher probability to find a buyer that is willing to pay more.

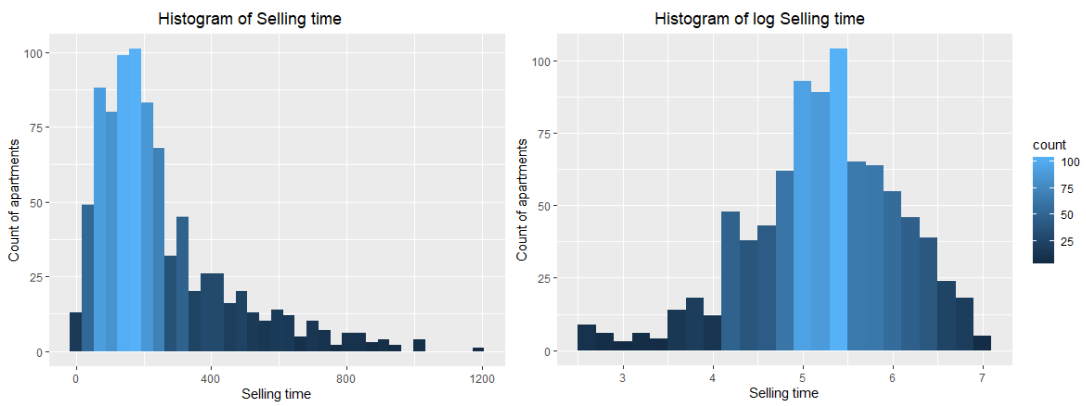


Figure 24: Histogram of TOM without transformation and after taking the logarithm

Different variables of the competition are selected for the analysis. The variable *competition* contains information of the amount of new competitive apartments launched to sale in nearby area. More specifically, it is the number of new apartments launched to markets in each

quarter by competitors in BA II area. These apartments belong only to residential projects. Most of these competitor projects are located 2-5 kilometers away from the project measured by the walking distance using Google Maps. Several projects are further than five kilometers away. Only 2 of the projects are located under 1-kilometer range. Ooi & Thao (2011) uses a similar variable that is used in this analysis, where they measure the change in housing stock as the new units launched by developers in each quarter.

The variable *Competition* has very high variation since its minimum value is 0 and the maximum value is 1311. In order to assess more reliably the impact of competition, the variable is divided into four groups so that numbers 1-4 describe the intensity of competition. If the new competition is considered low, the variable gets 1, and when competition intensity is fierce, the variable gets 4. Furthermore, dummy variables are created based on this logic as shown in Table 4. The groups were divided so that each group has values and this is why equally growing groups does not make sense to use. For instance, there are no observations between 400-500 apartments.

Table 4: Dummy variable creation for describing the competition

Number of new apartments launched to markets by competitors	Dummy variables *C1, C2, C3* and *C4* for each competition category
0-99	*C1,* if variable belong to group it gets value = 1, if not = 0.
100-199	*C2,* if variable belong to group it gets value = 1, if not = 0.
200-499	*C3,* if variable belong to group it gets value = 1, if not = 0.
500 or more	*C4,* if variable belong to group it gets value = 1, if not = 0.

A variable *CPrice* stands for the average price without tax in square meters of new apartments launched for sale by competitors in BA II area. This variable allows to see whether the competitor's strategy on prices affects the project sales. The prices of new launched apartments have been mostly growing, as presented in section 4.1.1. As shown in Table 5, the average price is 2298.

There are relatively high amount of newly-built apartments available in the markets, which differs from the variable of competition. The variable *Supply* describes the overall supply in the primary markets in BA II area. More specifically, it is the total number of newly-built available apartments in each quarter. The minimum number of available apartments is 312 and the maximum is 995. These apartments describe the overall situation well, but the apartments that have been in the market for a comparably long time are not expected to affect to sales of the project as much as the new competition does.

Economic fundamentals often affect to house prices, as well as TOM (Yu 2010; Gardner 1989). It is interesting to see how macroeconomic circumstances affect to project. Such as Kang & Gardner (1989) states that at times the mortgage rates are increasing, demand for housing decreases. As the financial costs rise, the buyers try to obtain price concessions to compensate for their "loss". Resulting from this, the seller might for instance, wait to obtain fair price, thus the TOM gets longer. Variable of interest rate is included to model to capture the market conditions. This variable *LR* stands for lending rate, which is the interest rate for household mortgages in Slovakia provided by National Bank of Slovakia (NBS 2022).

Competition variables as well as the lending rate are lagged. In selling price analysis, the value of one quarter prior sales is used and the selling time analysis uses the launch quarter value. This means that in selling price analysis for instance if an apartment is sold in the second quarter of 2020, the previous quarter's value will explain this observation. It is not feasible to use sale quarter values as an explanatory variable. The apartment might have been sold at the beginning of quarter, thus the future value shouldn't explain the variable in such vigorous intensity, although the future expectations of interest rates might have some explanatory power. For example, Ooi & Thao (2012) suggest that when supply increases, the existing prices responds in the next quarter.

Table 5: Descriptive statistics of the data

n = 865

variable	mean	sd	median	min	max	skew
SP	138838	43522	130863	66714	328416	1,22
TOM	253	199	194	14	1191	1,5
Rooms	2,31	0,8	2	1	4	0,15
Floor	4,92	2,11	5	2	12	0,63
Ground	0,31	0,46	0	0	1	0,82
Mid	0,48	0,5	0	0	1	0,07
High	0,17	0,38	0	0	1	1,75
Top	0,04	0,19	0	0	1	4,9
Size	58,8	16,39	51,8	36,5	116,9	0,6
Balcony_size	6,13	2,87	5,26	0	15,43	0,83
Terrace_size	2,07	9,36	0	0	85,98	5,47
Spring	0,3	0,46	0	0	1	0,86
Summer	0,36	0,48	0	0	1	0,57
Fall	0,34	0,47	0	0	1	0,7
Supply	732	161	757	312	995	-0,37
CPrice	2298	383	2304	1747	3279	1,13
LR	1,36	0,33	1,36	0,96	1,96	0,34

Finally, Table 6 introduces the correlations between variables. Examining the correlations is interesting to see the correlations between dependent variables price and TOM but also to detect possible multicollinearity in the data. Most of the apartment characteristics have a positive correlation with the price. The lending rate negatively correlates with price, whereas competitor price positively. Competition variables correlate positively and negatively with price. As expected, there is a high positive correlation between *Rooms* and *Size*. Variables seem to have low or moderate correlation with selling time. The numerical floor variable has the highest correlation with selling time (0.41). The dummy variable *mid* correlates also with selling time (0.27). The competition dummy variables have a positive correlation when competition is low and negative when high. Furthermore, the selling price correlates with selling time. Overall the variables have both negative and positive correlations with TOM.

Table 6: Correlation matrix

	SP	ST	Rooms	Ground	Mid	High	Top	Size	Balcony _size	Terrace _size	Spring	Summer	Fall	AVAP	CP_SP	C1	C2	C3	C4	LR
SP	1,000																			
ST	0,253	1,000																		
Rooms	0,753	0,051	1,000																	
Ground	-0,185	-0,223	-0,030	1,000																
Mid	-0,031	-0,067	-0,041	-0,648	1,000															
High	0,052	0,268	-0,051	-0,304	-0,437	1,000														
Top	0,432	0,191	0,283	-0,132	-0,189	-0,089	1,000													
Size	0,821	0,077	0,939	-0,048	-0,016	-0,061	0,281	1,000												
Balcony_size	0,078	-0,084	0,123	-0,099	0,193	0,027	-0,319	0,131	1,000											
Terrace_size	0,321	0,163	0,256	-0,024	-0,214	-0,056	0,735	0,211	-0,384	1,000										
Spring	0,056	-0,243	-0,011	0,014	0,019	-0,037	-0,009	-0,015	0,054	0,015	1,000									
Summer	-0,094	-0,012	-0,028	-0,017	-0,043	0,069	0,018	-0,034	-0,113	-0,014	-0,496	1,000								
Fall	0,042	0,249	0,039	0,004	0,025	-0,034	-0,009	0,049	0,063	0,000	-0,468	-0,535	1,000							
Supply	-0,033	-0,147	-0,061	-0,033	0,034	0,032	-0,071	-0,079	0,115	-0,042	0,303	-0,234	-0,056	1,000						
CPrice	0,425	0,189	-0,069	-0,103	-0,010	0,121	0,038	0,009	-0,029	-0,038	0,207	-0,155	-0,044	0,042	1,000					
C1	-0,028	0,108	0,002	0,054	-0,017	-0,058	0,027	0,016	-0,128	0,029	-0,214	0,105	0,102	-0,569	-0,083	1,000				
C2	0,201	0,090	0,012	-0,023	-0,020	0,029	0,051	0,041	0,007	0,025	0,011	-0,146	0,138	-0,087	0,168	-0,365	1,000			
C3	-0,250	-0,148	0,028	0,004	0,043	-0,042	-0,040	-0,027	0,070	-0,008	0,010	0,179	-0,192	0,137	-0,381	-0,334	-0,424	1,000		
C4	0,079	-0,053	-0,049	-0,036	-0,007	0,076	-0,044	-0,034	0,048	-0,053	0,208	-0,147	-0,053	0,562	0,331	-0,251	-0,319	-0,292	1,000	
LR	-0,491	-0,237	0,039	0,073	0,005	-0,076	-0,042	-0,061	0,043	0,057	-0,068	0,175	-0,112	-0,023	-0,791	0,020	-0,337	0,535	-0,239	1,000

4.2.2 Dividing data by type of an apartment

Data is divided into three subgroups because of two main reasons. At first, the variable for room type correlates highly with the size variable, as shown in the correlation matrix in section 4.2.1. Secondly, by examining each type of apartment group separately, one can see if competition affects these groups in different intensities. The data contains four different types of apartments, but since the share of 4-room apartments is relatively small (57 observations), the three- and four-room apartments are combined into one subgroup. These two types can be combined because both of types of apartments are quite large measured by area and are considered to be family apartments. These apartments have similar characteristics in terms of their sales and demand.

For subgroup analysis purposes, two new variables are presented. The first one is *Supply_type* and the second one *CPrice_type*. The first one contains information of available newly built apartments in residential projects of each type in BA II and latter the average prices for each type of apartment. For example, one-room apartments are explained using available 1-room apartments and their average prices. By doing this, the impact of competition is more realistic because it is assumed that for example, the studios do not compete against four-room apartments. The buyers most often have a preference for a certain size of apartments.

Creating the three- and four-room subgroup requires some processing of data. The amount of available three- and four-room apartments are combined to create a variable to measure supply of these apartment types. The average price of these apartments are calculated by taking the average between the three- and four-room apartment price without vat. Four-room apartment average prices have two missing values, and those are replaced by using the average price of the previous four quarters. These variables are on a quarterly basis and assigned for each observation by lagging the value based on the prior quarter of sale or the marketing quarter, similarly, as explained in chapter 4.2.1. The following Table 7 describes the three new subsets that are used to create separate models for each type of apartment.

Table 7: Summary statistics for subgroups

	1-room apartments / n=130			2-room apartments / n=396			3-and 4-room apartments / n=339		
	mean	**min**	**max**	**mean**	**min**	**max**	**mean**	**min**	**max**
Price	98223	66714	148626	120984	72522	192748	175270	108898	328416
Size	38,19	36,5	40,5	49,99	43,4	75,1	77,01	63,4	116,9
Ground	0,32	0	1	0,32	0	1	0,29	0	1
Mid	0,5	0	1	0,49	0	1	0,46	0	1
High	0,18	0	1	0,18	0	1	0,15	0	1
Top	-	-	-	0,01	0	1	0,09	0	1
Terrace_size	-	-	-	0,79	0	37,25	4,36	0	85,98
Balcony_size	4,93	2,6	6,35	6,54	0	9,88	6,11	0	15,43
TOM	294,6	15	1001	218,24	14	1191	277,71	17	1003
Spring	0,3	0	1	0,31	0	1	0,3	0	1
Summer	0,38	0	1	0,37	0	1	0,35	0	1
Fall	0,32	0	1	0,32	0	1	0,36	0	1
Supply_type	69,61	30,3	117,52	341,34	134,16	457,7	315,74	143,52	469,91
CPrice_type	109381	55188	151426	143389	112192	194736	240814	160084	355269
C1	0,24	0	1	0,22	0	1	0,22	0	1
C2	0,36	0	1	0,29	0	1	0,34	0	1
C3	0,19	0	1	0,31	0	1	0,28	0	1
C4	0,21	0	1	0,19	0	1	0,16	0	1
LR	1,29	0,96	1,9	1,39	0,96	1,96	1,37	0,96	1,96

The room variable is naturally eliminated. One-room apartments do not have any terraces or apartments on top floors, thus terrace size and top variables are not used. The summary shows that there are a lot of differences between subgroups. The one-room apartments' average price is 98 223 euros, 2-room apartments 120 984 euros and three- and 4-room apartments is 175 270 euros. The price deviation is the largest for three- and four-room subgroups, which is quite logical. The project and competitor prices are not fully comparable in this context because the project price does not include VAT but the competitor price does. However, the tax percentage is 20 % thus by calculating the tax-free price of competitor price it actually shows that on average, project price for one-room and two-room apartments has been higher, and price for three- and four-room apartments has been lower than competitor’s price. This comparison also shows that the average selling times are quite similar and there are only small differences between groups. The figures indicate that the smallest apartment takes more time to sell, and for all subgroups, there has been at least one project with an extremely long selling time which is more than 1000 days. Two-room apartments have a shorter mean and median selling time than other apartment types.

5 Results

The following chapters present the results for price and selling time models. At first, the coefficient transformation logic is presented. Then the results of an overall price model for whole data set is introduced, after which the models and their results for each type of apartments are discussed. This is followed by the results of TOM models so that at first, the result for the overall TOM model is presented after which the data is divided by the type on an apartment, as done in price analysis. The chapter ends with a discussion of the model specification.

5.1 Coefficient transformation

Interpretation of logged variables is slightly different than in basic linear model. The final regression models contains both logged dependent and independent variables. If the dependent variable is log-transformed and the explanatory variable is not, the coefficient is exponentiated as follows: $(e^{coefficient}) – 1) * 100$. This coefficient is interpreted so that an unit increase in the explanatory variable is associated with this percentual change in the dependent variable. If both the dependent variable and independent variables are log-transformed, one percent increase in the explanatory variable is associated with $(1.01^{coefficient} – 1) * 100$ percent change in the dependent variable. (Yang 2020) The coefficients of models that uses log transformations in this study are transformed using these equations. Linear model coefficients can be used interpreted as is.

5.2 Selling price model

This section presents the regression results of the overall price model using the whole data set. This section aims to answer the research sub-question and to analyse which factors have an impact on the final sale price. The special interest is to figure out whether competition affects the apartment prices of the project. The final model is constructed based on the results

of the specification tests as introduced in chapter 5.4. Outlier detection resulted in removal of 4 observations.

As already presented, there exists interrelationships between explanatory variables, which indicates a presence of multicollinearity according to Singla and Priyanka (2019). The VIF values are calculated for the price model to check that the multicollinearity problem does not exist. If the VIF value exceeds the threshold of 10, a multicollinearity problem is present (Montgomery, Peck & Vining 2012).

As suggested by Dormann et al. (2013) to remove multicollinearity from the model, the correlating variables, proven by high VIF values are eliminated. The dummy variables *Balcony* and *Terrace* are eliminated because of multicollinearity, as discussed in section 4.2.1. The VIF values for the final overall price model are calculated in Table 8 in which the variable *Rooms* have quite high VIF value (9.42) with the *Size* (9.38). This is not a surprising result because as number of rooms increase, most likely the area of an apartment gets larger. A dummy variable for each room are tested if the multicollinearity can be reduced. Unpredictably, the change to dummy variables does not solve the multicollinearity problem, instead it increases the VIF values for *Room2*, *Room4* and *Size* to exceed the threshold. Considering this, *Rooms* variable is left to analysis as numerical variable. The variables for rooms and size are kept in the model because they are considered to be important factors explaining price. However, the division to subgroups reduces the problem, and the results are discussed the following chapter 5.2.1.

Table 8: Variance inflation factors (VIF) of independent variables for overall price model

Variable	VIF	Variable	VIF
Rooms	9.42	Summer	1.73
Ground	1.20	Fall	1.74
High	1.24	AVAP	2.42
Top	2.42	CP_SP	3.30
Balcony_size	1.34	C2	2.19
Terrace_size	2.46	C3	2.70
Size	9.38	C4	3.54
ST	1.37	LR	3.73

Log-linear model is used for the overall price model. After taking natural logarithmic of the price, the model for overall price model becomes:

$$\ln(P) = \beta_0 + \beta_1 Rooms + \beta_2 Size + \beta_3 Ground + \beta_4 High + \beta_5 Top + \beta_6 BalconySize + \beta_7 TerraceSize + \beta_8 TOM + \beta_9 Summer + \beta_{10} Fall + \beta_{11} Supply + \beta_{12} CPrice + \beta_{13} C2 + \beta_{14} C3 + \beta_{15} C4 + \beta_{16} LR$$

where *Supply* describes available newly-built apartments in BA II, *CPrice* stands for competitor price and C2-C4 are dummy variables and describe the competition in the area, where C1 stands for low and C4 high competitive intensity. The variable *LR* presents the interest rate for mortgages. Market-related variables *Supply*, *CPrice*, *C2*, *C3*, *C4* and *LR* are lagged to one quarter before the sale of a unit. Consequently, a transaction is explained using the value of one quarter before the sale occurs.

The results of the overall price model are presented in Table 9. The explanatory power of the model shows that the model explanatory power is rather good, and the coefficients are significant at a general level. Only one factor, a dummy variable for the competition (C2) is not significantly explaining the price, if using the 95 % confidence interval. However, it is important to note that the overall price model has heteroskedasticity, thus the significance of parameter should not be interpreted using the usual 5 % significance level. This is why 0.01 % significance level is presented in this and all of the following models, to make a more reliable interpretation of variable significance. The regression outcome present one level as reference value in the case of categorical variables which are the floor and competition (C1-C4). Mid floor is selected as a reference level among the floor levels Ground, Mid, High and Top. C1 is selected as a reference among the competition levels C1, C2, C3, C4. This means, when interpreting the results, one should compare the results to the reference level. For example, if the coefficient for Ground floor is positive and statistically significant in the price model, it can be stated the model predicts that the apartment in ground floor has higher prices than in mid floors.

Table 9: Overall price model results for whole data set

	Estimate		Std. Error	t value	p value
(Intercept)	11,19212	***	0,04259	262,815	2E-16
Rooms	12,211	***	0,00773	14,905	2E-16
Size	0,827335	***	0,00038	21,731	2E-16
Ground	-5,33699	***	0,00476	-11,517	2E-16
Mid	ref		ref	ref	ref
High	3,149826	***	0,00598	5,182	2,74E-07
Top	9,216733	***	0,016823	5,24	2,03E-07
Balcony Size	0,247203	**	0,00081	3,028	0,00254
Terrace Size	0,262352	***	0,00034	7,723	3,22E-14
TOM	0,005637	***	1,19E-05	4,725	2,7E-06
Summer	-1,34116	*	0,00551	-2,448	0,014558
Fall	-2,19581	***	0,00564	-3,938	8,9E-05
Supply	0,005447	**	1,95E-05	2,795	0,005303
CPrice	0,011583	***	9,56E-06	12,116	2E-16
C1	ref		ref	ref	ref
C2	0,200568		0,00643	0,312	0,755229
C3	-1,57856	*	0,00737	-2,16	0,031059
C4	-3,22652	**	0,00984	-3,332	0,000898
LR	-28,4159	***	0,01191	-28,05	2E-16
R^2	0.96				

* Indicates significance at 5 % level

** Indicates significance at 1 % level

*** Indicates significance at .01 % level

Competition: The relationship between the price and competition is mixed to some extent. The coefficient for the competition dummy C2 is positive but not significant, whereas higher competition gets negative coefficients. C1, the low competitive intensity is used as a benchmark and each dummy variable is compared with this reference group. The results indicate that at times of relatively high competitive intensity (C3), the price decrease by 1.58 % when comparing to the reference. The price drops even more, by 3.2 % when the competitive intensity is fierce (C4). This result generalizes the findings of previous studies such as Hui et al. 2016 (2020) and Bulan et al. (2009), who found that competition decreases the price. The competition levels C3 and C4 are only significant at 5 % and 1 % significance levels, thus it is not safe to say competition has statistically significant relationship with price. However, it seems that high competition intensity has some negative impact on price. One of the reasons for negative coefficients can be a price-sensitive seller when number of competitors rise.

Supply: Surprisingly, the supply has a positive relationship with price, but again, the variable is significant only at 1 % significance level, thus not significant at 0.01 % level. The result suggests that a one-unit increase in supply increases the price by 0.005 %. The result disputes against the hypothesis that suggest the increase in supply would decrease the price. However, this observation can be explained for example with increased market activity at times when supply is increasing. Then the market conditions might be favourable and demand rising, thus company can yield higher prices.

CPrice: Competitor price describes the average price of newly-built apartments that are launched to sale. The relationship between the project and competitor price is positive. This is relatively a good indicator, that the project price is following the area trend. When competitor increases their prices, the project prices tend to increase as well.

Rooms: The results suggest the more rooms the apartment has, the higher the price is. This is very reasonable, and the variable is also statistically significant. The model suggests 12.2 % increase as the room numbers of an apartment increase by one.

Size: As an apartment gets larger, the higher the price will be. The variable is statistically significant, and the coefficient is 0.8 %, which means one square meter increase in area, increases the price by 0.8 %.

Floor: Ground, High and Top variables represent the floor level of an apartment. The model finds significant effect on each floor and the results are as one can expect. The apartments on higher floors are more expensive. This is because there is more demand for apartments located in higher floors. The Ground floor has negative coefficient because Mid floor is used as a reference. This means, the price in lowest floors is approximately 5.3 % lower than in the mid floors. Previous literature unequivocally supports the fact that houses located on higher levels get higher premiums. The observation seems universal (see e.g., Baranzini & Ramirez 2005; Wong et al 2011; Ong & Koh 2000).

Balcony and Terrace size: The area of a terrace both have positive and significant impact on price. The coefficient for balcony is positive but significant only at 1 % significance level.

TOM: Increase in selling time increases the price. The impact is small, but statistically significant, and one day increase in selling time increases the price by 0.0056 %. This means that when the apartment marketing time gets longer, there is a small effect on price. Although

the theory has evidence for and against, most of the research supports this finding that as TOM increases, the final price rises (see e.g., Yavas 1992).

Season: The season variables Summer and Fall also impact prices, but summer only at 5 % significance level. Spring is the benchmark thus the negative coefficients for summer and fall indicate that the company manages to yield higher prices in Spring. This means there is a slight price decline for apartments sold in Summer and Sall. From these three categories, the prices are lowest in Fall.

Lending rate: The variable *LR* stands for the interest rate for mortgages. The model predicts that the lending rate is significant and powerful factor in explaining the price. The results suggest that a one-unit increase in interest rate decreases the price by 28 %.

5.2.1 Selling price model results by type on an apartment

This section aims to determine whether there are differences in the variable effects on different types of apartments. Separate price models are estimated for each subgroup. The data is divided based on the apartment type, as presented in chapter 4.2.2. The selected models are all log-linear models. At first, the VIF values are calculated for each subgroup model in Table 10 to test whether multicollinearity still exists. It can be concluded that the models do not detect problematic multicollinearity.

Table 10: Variance inflation factor (VIF) values for each subgroup

	1-room	2-rooms	3-to 4-rooms
Variable	VIF	VIF	VIF
Size	1,78	1,24	1,67
Ground	1,30	1,38	1,43
High	1,30	1,26	1,42
Top	N/A	1,20	3,54
Balcony_size	1,44	2,09	1,70
Terrace_size	N/A	1,56	3,26
TOM	1,87	1,52	1,37
Summer	1,67	2,00	1,61
Fall	2,10	1,89	1,77
Supply_type	3,02	3,75	1,37
CPrice_type	3,93	3,84	6,56
C2	2,53	2,67	1,93
C3	2,81	4,04	2,73
C4	4,55	4,71	1,94
LR	5,54	4,67	5,85

The model for subgroups differs from the overall price model regarding supply and competitor price variables. Each apartment type is explained using the supply of similar type apartments in BA II area. Similarly, the competitor price here refers to the average price of similar type of available apartments. The competition dummy variables are the same as the overall price model used. Putting all these variables together, the model for the one-room apartments is as follows:

$$\begin{aligned}\ln(P) = \ & \beta_0 + \beta_1 Size + \beta_2 Ground + \beta_3 High + \beta_4 BalconySize + \beta_5 TOM \\ & + \beta_6 Summer + \beta_7 Fall + \beta_8 Supply\_type + \beta_9 CPrice\_type \\ & + \beta_{10} C2 + \beta_{11} C3 + \beta_{12} C4 + \beta_{13} LR\end{aligned}$$

where the variable *Supply_type* is available one-room apartments one quarter before a sale, *CPrice_type* is the average price of one-room available apartments, and *C2, C3,* and *C4* are dummy variables describing the overall competition similarly as used before. These variables follow the same logic in two-room and three- to four-room models. For example, *Supply_type* in the two-room model is the available two-room apartments and in three- to four-room model it is the available three- to four-room apartments. Models for two- and three- to four-room apartments are estimated using the following equation:

$$\ln(P) = \beta_0 + \beta_1 Size + \beta_2 Ground + \beta_3 High + \beta_4 Top + \beta_5 BalconySize + \beta_6 TerraceSize + \beta_7 TOM + \beta_8 Summer + \beta_9 Fall + \beta_{10} Supply\_type + \beta_{11} CPrice\_type + \beta_{12} C2 + \beta_{13} C3 + \beta_{14} C4 + \beta_{15} LR$$

The results for each subgroup model are presented in Table 11, which presents the coefficient and standard error for each variable in each model. In terms of R-square, the fit for each model is favorable and roughly the same, but lower than in the overall price model. This is mainly because these subsets are more homogenous, and contain less variance, hence the model can not explain that much of the variance in the dependent variable. Surprisingly, the R-square is good for one-room model although it has smaller number of observations, but more variables are significant for larger apartments. This is mainly because they are more unique in their attributes when compared to one-room apartments. The price models have some heteroskedasticity, thus the 0.01 % significance level is used.

Table 11: Price models for each subgroup

	1-room model		2-room model		3- & 4-room model	
	Estimate	Std.Error	Estimate	Std.Error	Estimate	Std.Error
(Intercept)	10,9171 ***	0,223	11,0317 ***	0,064	11,0552 ***	0,083
Size	1,9014 ***	0,004	1,5400 ***	0,001	1,0884 ***	0,000
Ground	-4,3776 ***	0,011	-5,0490 ***	0,006	4,2672 ***	0,008
Mid	ref		ref		ref	
High	6,2683 ***	0,013	2,6872 ***	0,007	7,8240 ***	0,010
Top			7,9634 *	0,036	21,1907 ***	0,021
Balcony_size	-1,9102 ***	0,004	0,5481 **	0,001	0,5111 ***	0,001
Terrace_size			0,6129 ***	0,001	0,2440 ***	0,000
TOM	0,0058 *	0,000	0,0102 ***	0,000	0,0102 ***	0,000
Spring	ref		ref		ref	
Summer	-1,2002	0,012	-3,8171 ***	0,007	0,1154	0,008
Fall	0,8481	0,014	-2,4918 **	0,007	-2,9385 **	0,009
Supply_type	0,0577 *	0,000	0,0138 **	0,000	-0,0014	0,000
CPrice_type	0,0003 ***	0,000	0,0002 ***	0,000	0,0001 ***	0,000
C1	ref		ref		ref	
C2	3,2802 *	0,015	-3,0557 **	0,008	4,8271 ***	0,009
C3	4,2092 *	0,019	-2,2835 *	0,010	5,1500 ***	0,012
C4	2,9325	0,023	-4,0573 **	0,013	2,2307	0,012
LR	-27,6942 ***	0,035	-28,2598 ***	0,015	-20,8309 ***	0,023
N	129		368		338	
R^2	0.94		0.94		0.94	

* Indicates significance at 5 % level

** Indicates significance at 1 % level

*** Indicates significance at 0.01 % level

Competition: Competition effect is different for each type of apartments. There seems to be a positive relationship between competition and price for one-room apartments, but only C2 and C3 are statistically significant at 5 % significance level. The reason why competition does not have such strong impact on prices of studios might be that the investors are usually interested in the one-room apartments which increases their demand. The results indicate that increased competition for three- to four-room apartments increases their prices, but the intense competition (C4) does not affect impact on the price. This result is against the theory that suggest increased competition decreases the prices. It might be that intense competition is a good sign to markets and it makes the area more attractive for the byers, which is why the C4 does not have a statistically significant relationship with price. To be more precise, increased building activity can mean the area is developing.

The 2-room model shows opposite results, where an increase in competition tends to decrease the price, but the variables are not significant at 0.01 % level. However, these competition dummy variables are very close to significance level of 0.01 %. The result for two-room apartments can be due to the high number of two-room apartments produced, as shown in the section 4.1.1, Figure 17. Most of the apartments produced are two-room apartments, whereas smaller amount of one-, three- and four-room apartments are produced.

Supply: The supply coefficients are mixed with the results of the competition, similarly as in the overall price model. One- and two-room apartment models suggest an increase in price while supply increases whereas three- to four-room model propose a decrease in price. However, only two-room apartment supply is significant at 1 % significance level whereas the one-room at 5 % significance level. The contradictory effect of competition versus supply for two-room apartments can be reasoned because it is assumed that new projects that launch apartments on sale have a more powerful effect on price. In this case the effect is negative. However, the existing new supply does not press prices down, maybe because the case project apartments are desirable for the buyers or the increased supply acts as a positive market prospect. Previous literature has emphasized that the relationship between house prices and supply is not straightforward (Ooi & Thao 2011, 1447).

Competitor price: Competitor average price effect is similar for each model. The coefficients are significant and the proposed increase in project price as competitor price increase by one unit is 0,001-0,003 % depending on the model.

Size: The larger the area of an apartment, the higher the price is. The variable is significant in each model.

Floor: Addition to floor level seems to increase the prices and the variables are statistically significant. One- and two-room apartment models suggest that the prices are lowest on the lowest floors and the prices are more expensive on the highest floors. The largest apartment model proposes that the prices are lowest on mid floors and highest on top floors. The higher prices on bottom floors than mid floors can be for example because families might prefer larger apartments, and the families might value the easy access on the lower floors. This can be also due to other characteristics of an apartment that are not included in the model.

Balcony and Terrace size: Balcony size negatively affects house price in the case of one-room apartments. This can be explained by the fact that the balcony size does not grow along the area of an apartment. Although the largest studios have the largest balconies, many of the large studios still have very small balconies, whereas small studios have large balconies. This presses the coefficient to negative. The coefficients are positive for two-room and larger apartments.

TOM: Increase in selling time increases the price of each type of apartment.

Season: The launch season affects differently to final prices of different types of apartments. The variables for two-room model are significant and the result suggests, the two-room apartments sold in Spring have the highest prices and in summer the lowest. The sold season does not have an impact on the studios, based on p-value. It seems that the three- and four-room apartments sold in fall have lower final prices than those sold in spring.

Lending rate: Lending rate again has high and significant impact on prices and the results are very similar for each subgroups. One percent increase in interest rate decreases the price more than 20 %. The lending rate for two-room apartments is the largest.

5.3 Selling time model

This chapter presents the analysis and results of the overall TOM model. A regression model is performed to answer the sub-research question 1.2 that aims to find which factors affect TOM. The final model constructed is selected based on the results of basic linear regression

assumptions tests performed, Ramsey RESET test, and a comparison of R-square values. The selling time model takes advantage of log-log model, because of many reasons. At first, it has been widely used in previous research and secondly, by changing the TOM variable to logarithmic, R-square gets better, and lastly, based on RESET test, the model contains less misspecification when compared to linear and log-linear ones. Possible outliers are detected are two most prominent outliers are removed.

Selling time model uses competition variables using the launch quarter of each apartment. This is done, because it is most likely that the market conditions at the time the apartment has launched to sale have an impact on the sales speed. Most of the apartments are launched to sale closer to the end of the quarter. For example, if considering a situation where the competition is moderate, one can argue that the apartment launched to market at these times may accelerate sales, when comparing to time with a high competitive intensity. The selection to use the marketing quarter can also be justified by testing different lags of competition variables. Both one month prior to sale and marketing quarter are tested, and the marketing quarter showed significant improvement in the model each time, in terms of R-squared.

The initial selling time model has a serious multicollinearity problem. Eliminating the *Supply* factor and *LR* reduces the problem significantly, thus they are eliminated. The *Size*, *Rooms*, and *SP* are also correlating with each other. Using dummy variables instead of numerical *Rooms* variable does not reduce the multicollinearity proven by VIF values. Rooms variable is tested to drop out from the analysis, but it did only a small change to the model. On contrast, dropping variables selling price and size, does more harm than good, because they are important factors determining the selling time. Considering this, the variables are left to the overall analysis. The final model VIF values are presented in Table 12. To tackle the multicollinearity problem, different type apartments will be analysed separately in chapter 5.3.1 in order to eliminate the serious multicollinearity problem caused by variables *Rooms*, *Size* and *SP*.

Table 12: Variance inflation factor (VIF) values for overall TOM model

Variable	VIF	Variable	VIF
Rooms	12,72	ln(Size)	17,40
Ground	1,33	ln(CPrice)	4,28
High	1,25	Summer	1,58
Top	2,50	Fall	4,00
Balcony_size	1,44	C2	4,30
Terrace_size	2,86	C3	3,63
ln(SP)	13,87	C4	2,43

The final selling time model is log-log model, where the dependent variable selling time and *Size*, *SP* and *CPrice* are logarithmic. A logarithm of balcony size is not reasonable since it contains zeros, and also logarithm of dummy variables won't improve the model. The logarithm of Rooms reduces slightly the p-value of Ramsey's RESET test, thus it is left as is. The overall selling time model is defined by using the following function:

$$\begin{aligned}\ln(ST) = {} & \beta_0 + \beta_1 Rooms + \beta_2 Ground + \beta_3 High + \beta_4 Top + \beta_5 BalconySize \\ & + \beta_6 TerraceSize + \beta_7 \ln(Size) + \beta_8 \ln(SP) + \beta_9 Summer \\ & + \beta_{10} Fall + \beta_{11} \ln(CPrice) + \beta_{12}\mathrm{C2} + \beta_{13}\mathrm{C3} + \beta_{14}\mathrm{C4}\end{aligned}$$

where C2, C3, and C4 are the competition variables as used before. It is reasonable to assume, that as competition increases, the TOM will get longer. The overall TOM model results are summarized in Table 13. The explanatory power of 0.4 is much lower than in the price models. The p-values for most of the variables are significant at a 5 % significance level. Almost half of the variables indicate statistical significance at the 0.01 % significance level. The large difference in the variable coefficients can be explained by the size of these variables, meaning the small and large variables have different marginal effect on TOM as the variable increase by one-unit or one percent.

Table 13: TOM model results for whole data set

	Estimate		Std.Error	t value	Pr(>|t|)
Intercept	-1,2086		2,867	-0,422	0,673
Rooms	-30,906	**	0,099	-3,736	0,00020
Ground	0,442		0,056	0,079	0,93684
Mid	ref		ref	ref	ref
High	30,702	***	0,066	4,043	5,8E-05
Top	-3,508		0,190	-0,188	8,5E-01
Balcony_size	-4,065	***	0,009	-4,439	1,0E-05
Terrace_size	-1,360	**	0,004	-3,196	1,4E-03
log(SP)	3,596	***	0,283	12,55	< 2e-16
log(Size)	-1,723	***	0,342	-5,107	4,0E-07
log(CPrice)	-3,532	***	0,323	-11,184	< 2e-16
Spring	ref		ref	ref	ref
Summer	35,107	**	0,093	3,228	1,3E-03
Fall	92,640	***	0,060	11,012	< 2e-16
C1	ref		ref	ref	ref
C2	12,092		0,109	1,051	2,9E-01
C3	38,450	**	0,088	3,693	2,4E-04
C4	21,325	*	0,090	2,143	3,2E-02
R^2	0.40				
N	863				

* Indicates significance at 5 % level

** Indicates significance at 1 % level

*** Indicates significance at .01 % level

Competition: High and intense competition (*C3* and *C4*) impacts TOM, whereas the lower competitive intensity does not have an effect. However, C3 and C4 are only significant at 1 % and 5 % significance levels, respectively. The coefficient for *C3* suggest that there is a 38.5 % increase in TOM when the competitive situation intensifies when compared to reference level C1. However, the effect of fierce competitive intensity (C4) is not as large as the C3 level. This can mean, the high intensity (C3) captures the effect of increased competition and the TOM does not get longer when shifting to C4 level. In general, this result would mean that as the competitive situation gets more intense, there is some prolonging effect on the TOM, but the impact is not linearly growing. This result is in line

with the previous research that suggests that when competition increases, the TOM increases for a seller (Anglin et al. 2003).

Competitor price: Competitor price is negatively affecting project selling times. As the competitor price grows by one percent, the selling times decrease by 3.5 %. This can indicate that the project apartments become more attractive for the buyer if competitors raise their prices. This can also signify that when prices are higher, the markets are active and selling times get lower.

Rooms: The coefficient for variable room type indicates that the more rooms the apartment has, the shorter the TOM is. The studios take more time to sell compared to larger ones. This can indicate that there is less demand for one-room apartments in this particular area or the buyers prefers larger apartments. However, this relationship might not be that straightforward because based on the analysis made in chapter 4.1., two- and three-room apartments sell fast, but the four-room apartments most often have longer TOM. The results indicate that the larger apartments may be affordable and reasonable size for potential buyers. Especially the two-room apartments are in the lower price category when compared to three-room and larger apartments, thus are affordable for many buyers.

Size: The size has significant negative impact on TOM, which indicates the larger apartments sell faster. This observation is in line with previously discussed variable *Rooms*. Li (2004) has found a significant negative relationship between size and TOM, whereas both Kang & Gardner (1989) and Asabere and Huffman (1992) arguments that size is not explaining TOM.

Floor: It's interesting to note that only the *High* floor category has significant effect on TOM whereas top and ground floors are not explaining selling time. The coefficient indicates that high floor apartments have approximately 32 % longer selling times than the mid floors. The reason for this can be that apartments on high floors are most often large apartments which selling prices are high. The apartments on the top floor have higher demand because of general preferences of top floor view and privacy and the ground floors are cheaper and thus more affordable for a larger group of potential buyers. This might create the effect that TOM is higher on high floors. Previous literature does not unanimously support the finding. Li (2004) found that higher floor apartment has shorter TOM. In contrast, Ong and Koh (2000)

results indicate that units on lower and higher floors both has higher TOM. They explained this observation by the fact that demand is higher for top floors, and so are the prices.

Balcony and terrace size: Both balcony and terrace size coefficients indicates that increase in area of a balcony or terrace lowers the expected sale time of that house. This shows the buyers may have preferences for apartments with larger balconies and terraces.

Selling price (SP): The model suggests that higher-priced apartments have longer TOM. This is in line with the housing market theory. Such as Knight (2002) points out that the higher price makes the TOM longer. This phenomenon is captured in multiple studies. However, most of the researchers examine the degree of overpricing and its relationship to TOM (Li 2004).

Season: The season variables have explanatory power on TOM, but only the Fall is significant at 0.01 % significance level. The results indicate that apartments listed in Fall have the higher selling time if comparing to Spring. Haurin (1988) has found similar results and suggests that season has an impact on TOM, and marketing time is shorter during summer and spring when compared to winter.

5.3.1 Selling time model by type of the apartment

This section aims to deepen the research on TOM in order to achieve more fundamental results and answer to the second sub-research question. This is not only to get comprehensive scope for the analysis and reduce the multicollinearity proved earlier but to compare different types of apartments to see if the competition effect varies among the datasets. Separate models are estimated for each subgroup based on the apartment type.

The final VIF values for each subgroup are presented in Table 14. Serious multicollinearity problem occurs in the model for two-room apartments. Only removal of season variables removes the problem, thus season is left out from the 2-room apartment analysis to get more reliable coefficients. Some other variables have relatively high VIF values as well, but none of them exceed the threshold of 10.

Table 14: Variance inflation factor (VIF) values for each subgroup

	1-room	2-room	3- & 4-room
Size	1,666	1,957	3,207
Ground	1,537	1,549	1,232
High	1,752	1,297	1,291
Top	N/A	1,104	4,038
Balcony_size	1,250	1,973	1,828
Terrace_size	N/A	1,773	3,534
SP	8,961	6,917	7,163
Supply_type	6,936	3,265	3,426
CPrice_type	4,997	3,012	6,610
Summer	7,039	N/A	3,798
Fall	4,521	N/A	3,819
C2	6,321	2,783	6,880
C3	8,504	7,015	6,979
C4	5,473	4,142	1,773

The one-room data set is the smallest and different from larger apartments when comparing results of specifications tests between these apartment types. This is most likely because the one-room data set contains the least variance and the apartments are alike in size and characteristics. The linear model performs better in explaining the TOM when compared to the log-linear model. The one-room linear model is presented as follows:

$$\begin{aligned} TOM = {} & \beta_0 + \beta_1 \text{Size} + \beta_2 Ground + \beta_3 High + \beta_4 BalconySize + \beta_5 \text{SP} \\ & + \beta_6 Summer + \beta_7 Fall + \beta_8 Supply\_type + \beta_9 CPrice\_type \\ & + \beta_{10} C2 + \beta_{11} C3 + \beta_{12} C4 \end{aligned}$$

where Supply_type and CPrice_type are available newly-built one-room apartments and their average prices in BA II area, respectively. The variables C2, C3 and C4 are the dummy variables for competition, as used in previous models. Two-room and three- to four-room models gets both log-linear form which are estimated using the following equation:

$$\begin{aligned} \ln(TOM) = {} & \beta_0 + \beta_1 \text{Size} + \beta_2 Ground + \beta_3 High + \beta_4 Top + \beta_5 BalconySize \\ & + \beta_6 TerraceSize + \beta_7 \text{SP} + \beta_8 Supply\_type + \beta_9 Cprice\_type \\ & + \beta_{10} C2 + \beta_{11} C3 + \beta_{12} C4 \end{aligned}$$

where the *Supply_type* stands for the corresponding number of available newly-built apartments for each type of room. Meaning, in two-room model, supply is the available two-

room apartments, whereas three- to four-room data set is explained using available three- and four-room apartments. The variable CPrice_type logic is the same.

The regression results for each subgroups of apartments are presented in Table 15. The models are constructed for one-room, two-room and three- to four-room apartments separately. The coefficients vary among different types of apartments. In comparison of subgroups, the one-room model has better fit in terms of R-square. Each model contains several variables that are not statistically significant. It is important to note, one-room model does not contain any heteroskedasticity and the 5 % significance level can be used. For other models, the 0.01 % level is preferred.

Table 15: TOM models for each subgroup

	1-room model		2-room model		3- & 4-room model	
	Estimate	Std.error	Estimate	Std.error	Estimate	Std.error
Intercept	632,070	330,8	9,347***	0,54	25,6262***	4,17
Size	-18,770 *	8,6	-7,487***	0,01	-4,266 ***	0,44
Ground	13,562	23,1	14,694	0,08	8,625	0,07
Mid	ref		ref		ref	
High	-2,211	29,9	18,317	0,09	-7,983	0,09
Top	-		-43,645	0,61	-50,412 **	0,19
Balcony_size	38,556 ***	7,4	-6,234**	0,02	-1,327	0,01
Terrace_size	-		-3,084***	0,01	-0,301	0,00
SP	0,012 ***	0,0	0,005***	0,00	4,699 ***	0,35
Summer	-209,142 ***	47,8	-		-23,020 *	0,12
Fall	-71,568	39,3	-		-3,293	0,11
Supply_type	-5,073 ***	0,9	-0,653***	0,00	0,006 ***	0,00
CPrice_type	-0,011 ***	0,0	-0,003***	0,00	-4,576 ***	0,37
C1	ref		ref		ref	
C2	114,925 *	48,5	117,907***	0,11	-55,184 ***	0,17
C3	192,792 **	56,8	209,875***	0,16	-67,724 ***	0,16
C4	72,557	50,5	244,182***	0,15	-42,10 ***	0,10
N	129		368		338	
R2	0.78		0.55		0.53	

* Indicates significance at 5 % level

** Indicates significance at 1 % level

*** Indicates significance at .01 % level

Competition: Competition seems to affect at different intensities to different types of apartments. For two-room and three- to four-room models, all competition coefficients are

statistically significant at a 0.01 % significance level. As the competition increase to C4, the TOM tends to increase for two-room apartments by 244 % and decrease for larger apartments by 42 %. Decreasing effect to TOM can be due to area development, that means as competition increase it attracts buyers because the area is developing. A high competition (*C3*) affects one-room apartments, but intense competition is not statistically significant. This can indicate that moderate and high competition already captures the effect of competition. Interpretation is simple for this linear model, as the competition increase for one-room model to C3, the TOM increase by 193 days. The significant positive impact on two-room apartment selling times can be due to fact that the two-room apartments are produced the most. It might be that the competitors launch a lot of two-room apartments to the markets, which is seen as increased TOM in this room type because the buyers will have a lot of choices.

Supply: Supply coefficients are statistically significant in each model. For smaller and two-room apartments the increase in supply decreases the TOM, whereas for larger apartments increase in supply increase TOM. The explanation why supply decrease studios selling time but competition increase, can be due to fact that most likely new competition causes a negative effect on sales, whereas the supply figures contains the overall supply of studios that should not have such as strong impact on one-room apartment sales than the competition has. The negative coefficient for supply can be a consequence of several things. At first, this might indicate of a signaling effect, because buyers might think as supply is higher, the confidence has increased in the housing markets. There could be also another explanation, if assuming the supply and expectations are connected with business cycles, supply may be catching the general economic situation which shortens TOM.

Competition price: Similarly, it seems that competitor price influences TOM. At times when competitor increases their prices, the project selling time decrease. There can be at least two reasons for that. At first, when the competitor rises their prices, it can indicate from positive prospects and might indicate of strong demand. When demand is high, the theory suggests that the selling times are shorter. On the other hand, if competitor increases their prices, the project apartments might come more attractive in terms of their price from the buyer's point of view.

Size: Size shows negative coefficient for each subgroup. Larger apartments measured by area have shorter TOM. The variance in size is the smallest in one-room apartments and it

seems that larger studios has shorter TOM. The one-room apartments sizes are between 37-41 m^2 thus, only a small increase in the size of the apartment significantly increases the attractiveness of the apartment, if the TOM is considered to be a measure of attractiveness.

Floor: The floor does not seem to have a significant impact on TOM. The top floor of three- and four-room apartments is only one that is significantly affecting TOM. It seems that three- to four-room apartments in higher floors sell faster. The two-room model floor variables are significant at 10 % significance level, thus it is concluded that the floor lever does explain TOM for smaller apartments. These results are similar to what some researchers have found in their studies. Such as Ong and Koh (2000) did not always find statistical significant relationships between floor and TOM. However, the top floor coefficients are mainly positive. This is because the higher demand for top floors increases their prices, which makes the selling time longer. The results of this thesis do not confirm this finding.

Balcony and Terrace size: Neither the balcony nor terrace size has a statistical impact on TOM in the case of three- to four-room apartments. An increase in size decreases the TOM for two-room apartments and increases for one-room apartments. This means studios with smaller balconies are more attractive. It is important to note that the smallest one-room apartments have relatively large balconies, thus the interpretation is not straightforward. This can indicate that the larger balconies increases the TOM, because apartment with large balconies are very small and are not so attractive for buyers.

Selling price: Selling price have statistically significant impact on TOM in each subgroup. The effects are positive and the most powerful impact can be seen for 3- to 4-room apartments where one unit increase in price increases the TOM by 4,7 %. For smaller apartments, the effect is less than 1 %.

Season: The both one-room and three- to four-room apartments sell fastest in summer when comparing to Spring. However, the Fall variables are statistically significant for these two apartment groups only at 10 % significance level and it is considered that the launch on Fall does not have an impact on TOM. This result is contradictory to the overall TOM model results. One additional model is performed for two-room apartment data, without the competition variables, and with season variables to check whether it can provide some explanations to support the interpretation. The results for additional two-room model in

Appendix 4 shows that two-room units sell fastest in Spring, and because this apartment type dominates the data, it provides an explanation for the coefficients.

5.4 Specification testing

The chapters 5.2 – 5.3.1 introduce the detection of possible multicollinearity in each model. It is observed that the selling price, rooms, and size are causing some problems. Multicollinearity problems are typical in hedonic pricing models (Laakso 1997). Significant multicollinearity is present in the TOM models, but it is tackled by removing some of the variables, as explained before. The process of finding the most suitable form for models is lengthy thus, it is not presented as a whole. The OLS estimates are BLUE if the assumptions are met. Next, the specification for each model are presented and discussed.

The residual normal distribution is examined visually using normal Q-Q plot as shown in Appendix 5. The ideal situation is that the residuals follow the dashed line. The overall price model follows the line quite well. The other models do not follow the line at the lowest and highest quantiles. Further, the histograms indicate that the residuals of most of the models are quite normally distributed (Appendix 6).

The Shapiro-Wilk test is performed for each model to verify normal distribution observations. The p-values for overall and 3- to 4-room price models suggest rejection of null hypothesis of normality using 95 % confidence level. Each one-room price model (p-value for log-linear model = 0.678) and log-linear and box-cox models of two-room models (p-value for log-linear and box-cox models 0.670 and 0.660, respectively) fulfils the criteria of normality. Selling time models show worse performance in terms of Shapiro-Wilk test, because only the linear one-room model passes the Shapiro-wilk test and is normally distributed.

Residuals versus fitted values plot can be used to detect non-linearity, heteroskedasticity and outliers. The plots for each price model are presented in Figure 25. The overall and two-room model shows that the points are quite well randomly spread around the zero. From the one- and three- to four-room models there can be seen some non-constant variance. Overall and one-room models red lines do not quite follow the horizontal line which may indicate non-linearity. The larger apartments red lines follow the horizontal line well, which indicates

there is not any discernible non-linear trend. The red line represents a lowell fit to the plot. Basically it is smoothing over the points to look for patterns in the residuals.

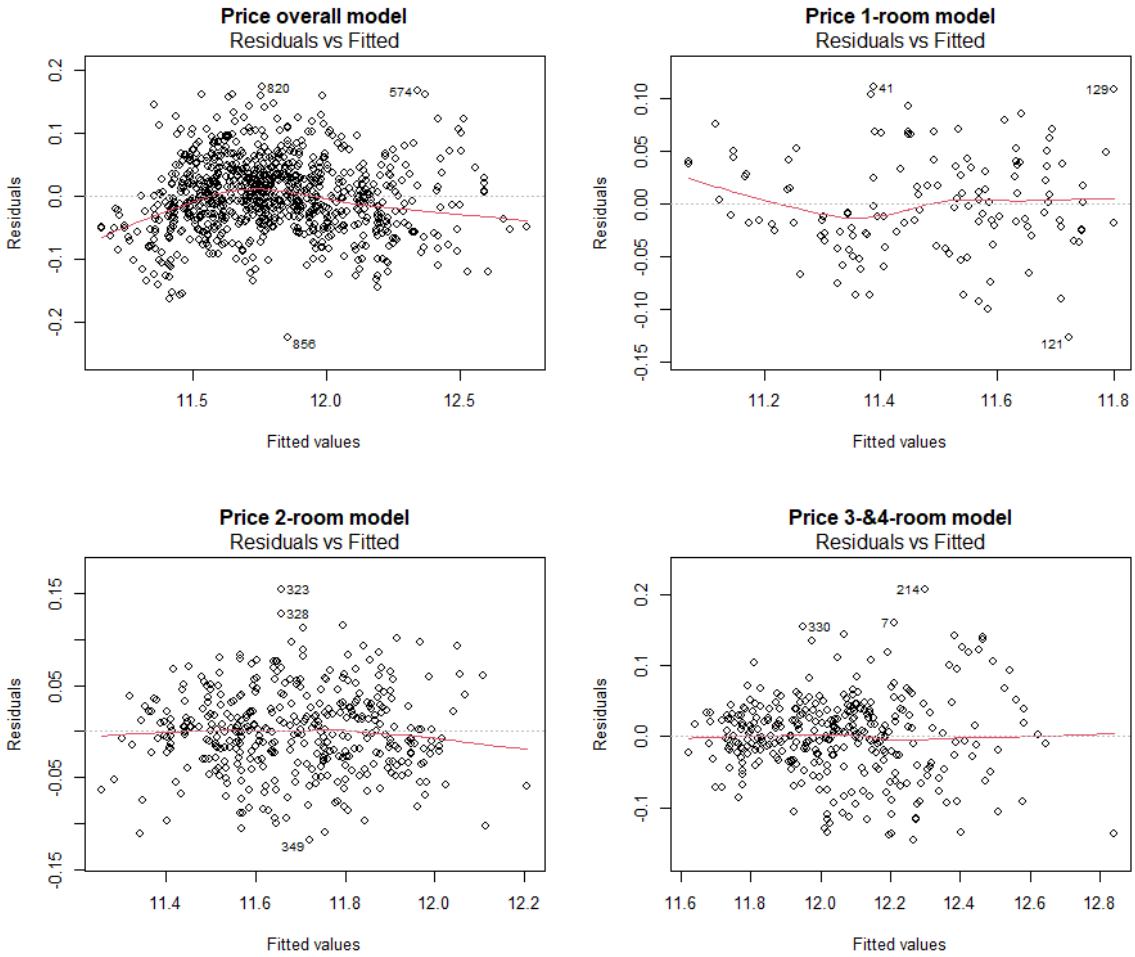


Figure 25: Residuals versus fitted values in price models

The residuals versus fitted plots for each TOM model are presented in the Figure 26. The overall model points are more centred slightly to left than randomly propagated around the line. Two-room model shows a decreasing trend in the residuals. The one-room model points are more centered on the left side whereas three- and four-room model points are propagated quite well around the zero line and these models do not indicate evidence of heteroskedasticity. The TOM models do not show such a good setup for linear regression compared to price models.

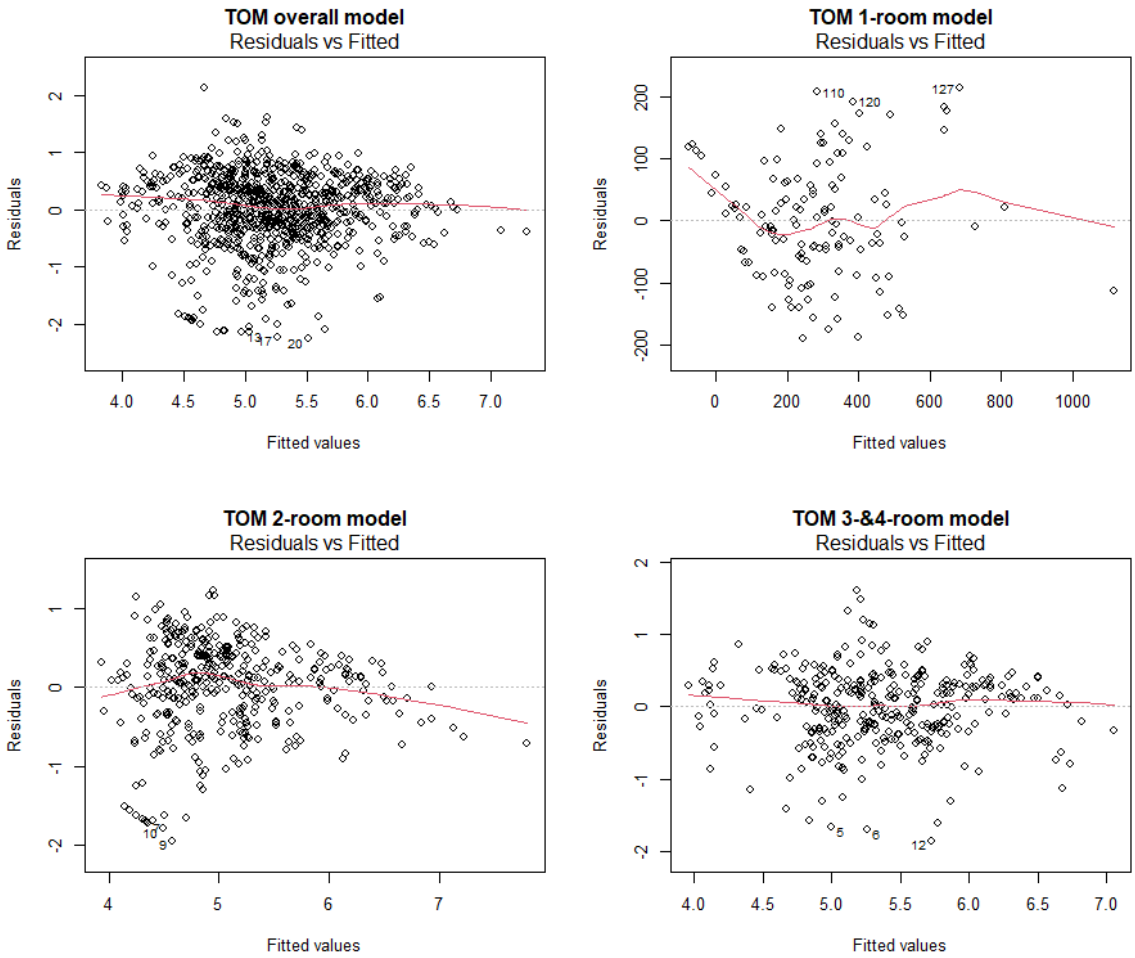


Figure 26: Residuals versus fitted values for TOM models

The models are tested for heteroskedasticity also using the Breusch-Pagan test as illustrated in Table 16. The smaller apartments seem to be more homogeneous, although the null hypothesis of homoskedasticity is failed to reject only at 1 % confidence level for price model. The p values for other models are below 0.05, thus the existence of heteroskedasticity cannot be denied. The larger apartments seem to have more complex variations in terms of their characteristics which the regression model struggles to handle, as suggested by Pakarinen (2018). Taking logarithmic transformation of explanatory variables could solve statistical problems; thus, the log transformations are tested for numerical variables Size and TOM. Performing these changes does not do much for the models. Log transformation of size increase the multicollinearity problem and transformation for explanatory variable TOM reduces the R-square. Considering these, it is decided to leave them as-is.

Table 16: Breusch-Pagan test

Breusch-Pagan test p-values

	Overall model	1-room	2-room	3-room
Price models	3.031e-14	0.0180	0.0004	3.613e-9
TOM models	2.2e-6	0.1292	1.254e-14	0.0050

Even though all of the assumptions are not met, the model performance has significantly improved by taking into use the logarithmic form in most of the models. The following residual versus fitted plot in Figure 27 shows the linear overall TOM model for comparison. It can be observed that the spread of residuals is increasing as fitted values change. This means, the model shows nonconstant variance which indicates presence of strong heteroskedasticity. The log-linear form for the overall TOM model fits much better than this linear one.

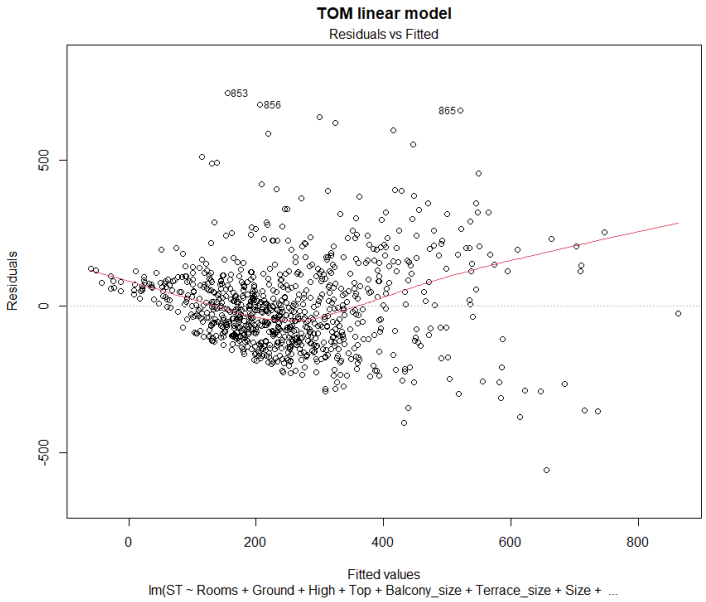


Figure 27: Heteroscedasticity in the linear TOM model

The model specifications are tested using Ramsey's RESET test. The null hypothesis is rejected if the model has incorrect functional form, error terms are correlated, or the model has omitted variables (Ramsey 1969). The RESET test results and R-square for each price model are presented in Table 17. The null hypothesis is not rejected for log-linear models for 2- and 3- to 4-room models, because the p-values are above the 5 % significance level.

1-room model RESET test p-value stays slightly under 0.05. Overall price model p-value gets close to 0, thus the one-room apartment model as well as the overall model contains some misspecification. The R-square can be used as an indicator of the explanatory power of the model (Li 2004, 317). The R-square values for each model are good.

Each log-linear price model is compared with the box-cox transformed model to see if they could provide better fit for the model. The box-cox transformation is tested for the response variable. The optimal lambda λ for the overall model is 0.1919. This indicates that the optimal transformation would be $(y^{\lambda} - 1) / \lambda$. Since the lambda is not far away from the zero, the logarithmic transformation could be suitable as well. The box-cox transformation is applied for price variable for each model, but it only improves the R-square slightly (0.9627) and based on Ramsey RESET test p-value (0.0000), it still does not fix the misspecification in the model, as shown in Appendix 7. Similar observation can be observed from subgroup price models. A quadratic model where the explanatory variables are in second power is also tested for the overall price model. It seems that quadratic and box-cox transformation provide only additional 0.0011 and 0.0012 increase in the model, respectively. Since the model already has very good explanatory power, it is decided to use the log-linear model because these alternative models would add complexity and do not significantly improve the model. The log-linear model has a more straightforward interpretation.

Table 17: Ramsey's RESET test and R-square for price models (log-linear)

Price Models

	Overall model	1-room model	2-room model	3- to 4-room model
R^2	0.9615	0.9354	0.9414	0.9373
Ramsey RESET test	0.0000	0.0435	0.6970	0.4306

Table 18 presents R-square and Ramsey RESET test results for TOM models. RESET test suggests that the model specification is well enough for most of the models. Overall model RESET test p-value is slightly above 0.05, whereas one-room and three- and four-room model p-values are significantly higher than 0.05. It can be concluded that the two-room model contains misspecification. Linear model showed the least misspecification and most significant explanatory power for one-room model, thus it was the main reason the linear model results are presented in this thesis. For other models, it is not so clear which models

are the best when comparing linear, log-linear and log-log models, as shown in Appendix 8. The log-linear model shows better model specification regarding Ramsey's test for three- to four-room model and well enough specification for two-room model. The RESET test p-values are poor for all one-room TOM models, which is not a surprise because of the availability of the data.

In terms of R^2 the explanatory power for each model is much lower than in price models. The one-room model manages to explain significantly more the variance in the dependent variable than the other models. The previous studies have mainly got low R-square values when using linear regression to estimate TOM. Such as Asabere and Huffman (1992) estimated TOM using linear regression to analyse residential sales data over four-year period. The R-square values of their models are between 0.16 and 0.25. In addition, Yavas & Yand (1995) used a two-stage regression model and got high R-square values for the price model, but the values for the TOM model remained below 0.20. Considering the previous literature, the R-square yielded in this study is at least above the mark.

Table 18: Ramsey's RESET test and R-square for TOM models

TOM Models

	Overall model (log-log)	1-room model (linear)	2-room model (log-linear)	3- to 4-room model (log-linear)
R^2	0.4001	0.7860	0.5448	0.5435
Ramsey RESET test	0.0885	0.2925	0.0001	0.1484

Based on the model comparison and residual analysis, most of the models have some statistical problems. The price model independent variables explain variance in the dependent variable much better than in the TOM models. In general, the subgroup models are less violating the basic assumptions of linear regression. It seems that smaller apartments are more homoscedastic. Heteroskedasticity problem can not be denied for most of the models.

Heteroskedasticity has been one of the major criticism of OLS models that has obtained especial attention in the previous research (see e.g. Stevenson 2004). It is worth noting the homoskedasticity assumption is required to justify the standard t and F tests, and confidence intervals for OLS estimation of the multiple linear regression model. As a result, if the heteroskedasticity is present, inferences might be distorted because standard errors as well

as F and t statistics can be invalid. Consequently, the significance of variables is stated as significant only if their p-values are significantly lower than the critical value in this thesis. This means, when the p-value of an variable is close to critical value of 0.05, it is safer to say the variable is not statistically significant, because the distribution might be distorted. Considering this, two orders of magnitude lower than the critical value of 1 %, that is 0.01 % is safer to use as a critical value to avoid mistakes when rejecting the null hypothesis for p-values. However, the estimates are not inconsistent or biased. According to Fletcher, Gallimore & Mangan (2000), when heteroscedasticity is present, OLS estimators are unbiased, but not the best. In conclusion, it turned out to be difficult to get rid of misspecification completely in this study. Considering this, there is some misspecification in the models and the coefficient should be interpreted judiciously.

6 Conclusions

This section aims to introduce and discuss the main findings of the study. The limitations for the research are discussed after which recommendation for future research is presented.

6.1 Main findings and discussion

Focusing on one multi-phase residential construction project located in Bratislava, the thesis examines which factors affect the final sale prices and time-on-the-market of the case project apartments. The main focus on this thesis is to figure out whether a competition plays a role in explaining the sales. The thesis aims to answer the following main research question: How to price apartments of a residential development project in different competitive market situations? Two sub-questions are used to answer to the main research question: Which factors have an impact on the final sale prices of a residential development project?, and Which factors explain the selling times (TOM) in a residential development project?

The hedonic pricing method is used as a framework in this study. Using the apartment transaction sales data, the actual number of new apartments launched by competitors, the number of available newly-built apartments overall in the nearby area, and other factors that can affect sales, hedonic regression models are employed. The study assumes that the buyers have a clear preference for newly-built apartments, thus this thesis considers the competition only in the primary residential real estate markets. The data is further divided into three subgroups to examine the effects of variables for each apartment type. Next, the main findings of sub-questions are presented and discussed. This section ends with a discussion to answer to the main research question.

The results indicate that multiple factors have an impact on the final sale prices and TOM. The findings also reveal that there are differences among the different type of apartments. These findings should be used as a support when pricing the apartments. The results suggests, the competition impacts more on TOM than the sale prices based on the variable significance. In addition, the overall supply of each type of apartment has a significant impact on selling times, but not for prices. It is interesting to note, that the launch season has

an impact on sales. In general, the results indicate that apartments listed in Fall have the highest marketing times. Previous literature has found similar results that marketing time tends to be shorter during Summer and Spring when comparing to winter (see e.g. Haurin 1988).

The subgroup analysis for each room type reveals that as competitive intensity increase, prices of large apartment tend to rise. However, the intense competitive intensity does not have an impact on price of large apartments. This finding shows evidence for contagion hypothesis, which predicts a positive relationship between supply and house price. However, the researchers have divided into two groups, where the other group supports competitive hypothesis and other one the contagion hypothesis. The competitive hypothesis suggests that as competition increase, the prices fall, as suggested by Ooi and Thao (2011). There does not seem to be impact of competition on prices of one-room apartments, but there can be observed some negative effect between competition and two-room apartment prices, although the effect is not statistically significant, but is very close to significance level.

In general, the previous literature provides support for the some of the findings and the results are inconsistent with economic theory. However, the literature do not provide support that the changes in competition would have an impact on the prices. Such as Hui et al. (2016) suggest that the price decreases for each marginal competitor located within 3 km radius and the price decreases even more when the competitor is located under 1km. This is because the apartments can be seen as substitutes for each other. To attract potential buyers, developers can lower their list prices.

Supply considers newly-built apartments of each type that are in sale whereas competition only the apartments of projects that are launched to sale in each quarter. From the point of view of a developer, one can expect that the competition would have stronger effect on the project sales. This means, the newly-built apartments that have been already for sale for a relatively long time, are competing with the project apartment, but should not have such a substantial impact on the sales than the new project apartments has. The results show some evidence for this assumption. In general, more competition dummy variables (C1-C4) seem be statistically significant than supply variables.

The findings also reveal that TOM is influenced by competition and supply, and the results vary between the apartment types. As competitive intensity increases, TOM gets longer for

one- and two-room apartments whereas an increase in supply shortens TOM for these apartment types. However, intense competition does not have an impact on one-room apartments. This kind of opposite effect of supply and competition variables was also seen in some of the price models. This makes it difficult to derive common conclusions of the effects of competition and supply to project sales. This can be evidence that supply is an indicator of market activity. When there is more supply, the markets are active, demand thrives, and indicates positive prospect for the buyers, creating a positive effect on price and negative on TOM. Nonetheless, if the new competition is intense, the price falls so that the company can maintain the market position, but also TOM can get longer.

Interestingly, the TOM gets shorter for three- to four-room apartments as competition intensity gets fierce. This can indicate from the fact that there are little four-room apartments available if comparing to other room types. As a result, the buyers don't have too much options available and sees project apartments desirable. Other explanation for this might be that the buyers sees increased competition as a positive sign, because when markets are active in the area, it indicates that the area is developing.

Multiple apartment characteristics and also interest rate impact on the sales of the project. Most notably, the findings reveal that interest rate for mortgages has significant negative relationship with price, that means as interest rate rises, the prices fall. Results also reveal that the prices are higher for apartments that have been on sale for longer time (TOM increases the price). Also, the prices are higher in the higher floors. The sale season has no or small impact on the final price. Two-room apartments have had the lowest prices when they have sold in Summer and highest when sold in Spring. Three-room apartments seem to have slightly lower prices if sold in Fall.

Results indicate that the larger apartments have shorter TOM, whereas more expensive apartments have longer TOM. Three- to four-room apartments located in the highest floors sell faster, but otherwise, there are no significant effects on floor level to TOM. In general, it seems that apartments that have launched on sale in Spring have a shorter TOM. In contrast, if launched on sale in Fall, has the longer TOM. However, when dividing the data into groups, it turns out that two-room apartments sell fastest when launched on sale in Spring, and one-room in Summer. The larger apartments launch schedule does not have an impact on TOM.

How should these results be used when pricing the apartments? At first, it is clear that multiple factors have an impact on the sale prices and TOM. Neither the competition nor supply of one-room apartments affects to this apartment type prices significantly. The competition has prolonging impact on TOM, but the high competition has no impact, and, the supply decreases the TOM, it seems that the price discounts are not necessary for one-room apartments even though the competitive level gets higher. Even small price premiums could be acceptable, if considering that there is a lot of investors buying studios that increases the demand.

As the number of large apartments in the markets rise, the three- and four-room apartment TOM gets longer. The same effect is not seen when the overall number of new competition that enters to markets, but it can be due to fact that most of competitor apartments are two-room apartments, and the actual number of larger apartments might not grow much, which is why the TOM does not get longer. Neither competition, nor supply negatively impacts on price, thus the competition is not necessarily a big threat to large units. Small price discounts can be profitable if a seller wants to speed up the sales.

As competition increase, it seems to affect heavily to two-room apartment selling times. When competition is fierce, the price discount would make the apartments more attractive if the seller wants to speed up the sales. However, the model did not find significant impact on the competition or supply to final sale prices, thus as the competition gets intense, it is not necessary to give price discounts to maintain profits.

As the results show the launch season has an impact on the TOM, the launch month should be considered when pricing the apartments. In general, Spring seems to be the best launch season in Bratislava and Fall the worst. To be more precise, the studios launched on sale in Summer sell fastest whereas two-room apartments in Spring. Larger apartment launch schedule does not have much impact on TOM.

The project dealt with in this thesis is a large-scale project and it sells the most desirable units at first. As time passes, the buyers will run out of time to buy the units they want. This means, as the supply of available units decreases in the project, the developer may use this as an advantage and raise prices for remaining apartments if the competitive intensity is not fierce.

6.2 Limitations

The models presented in this thesis do not pass all the statistical requirements, and more flexible methods are required to achieve more reliable results. In general, the problem of heteroskedasticity is a major criticism of the OLS model that has obtained in the literature. In multiple regression analysis, the homoskedasticity assumption has failed in many other research as well (see e.g. Stevenson 2004; Fletcher et al. 2000). When estimating the hedonic price function, it seems to be worthwhile to divide the data into more homogeneous subgroups and run separate regressions for them to reduce heteroskedasticity. These findings can be used as a benchmark for further research.

The R-square values for TOM models are moderate, and the models managed to explain the differences in selling times fine. Previous literature has found that TOM has significantly lower R^2 values in the hedonic pricing models. This can be because of many reasons, such as prices should capture both locational and physical differences of houses and these variables should not be significantly explaining TOM. Many different things could affect TOM, which is probably why its exploring is difficult. In many cases the power of TOM model is due to underpricing and overpricing. (Yavas & Yang 1995a, 365) This study does not take into account the over- and underpricing of apartments. Furthermore, one possible explanation to low R-square values can be that there is some other variables that could have influence on TOM, but might not be observable, such as seller's discount or interest rate expectations, as Yavas & Yang (1995a) points out. The variable selection in this study is restricted due to limited available data, thus this limitation can cause a bias in model estimation. The quality of the dataset is considered to be high, but the amount of information in it is not so comprehensive it could be.

When thinking the pricing strategies of a company, the time-on-the-market should not be unambiguously explaining the sales performance, which means the shorter TOM, the better. This is obvious because, as explained earlier, the apartments are launched to sale at an early stage. The developer may be willing to achieve longer TOM at the cost of waiting to gain more profits. This means, the lower TOM is not always better for the seller because the apartments are launched to sale at a very early stage before completion. However, the TOM is a good indicator of the most sought-after apartments because they are often sold quickly due to high demand. It is also important to note that the interpretation of the relationship

between TOM and price can be difficult because there can exists endogenous variable problem, meaning that price and TOM are jointly determined (Dubé & Legros 2015). DiPasquale (1999) also points out that the relationship between prices and supply is difficult to measure empirically.

Kang and Gardner (1989) have emphasized some implications about using market variables. They point out that the expected interest rates are not considered and the actual and expected macroeconomic variables can change between listing and sale date. Furthermore, the relationships in the model can vary among different times in terms of housing market conditions. This study uses data that contains one major economic crisis Covid 19, which can bias the results. However, Slovakia is facing housing scarcity, and Covid delayed some of the projects. This means, the Covid should not have negatively affected house sales, but the other way around. Bratislava would need even more construction so that the supply would meet the demand.

The assumption made in this thesis is that all of the launched apartments in the vicinity are competing with project apartments. In reality, not all of these apartments are substitutes and compete with each other. By selecting only the competing projects and apartments by similar decoration, special characteristics and building types might have yielded more insightful results, if would have been available. In addition to competition in primary markets, houses in the secondary market compete with newly-built houses to some extent. This study does not consider the competition in the secondary markets.

6.3 Future Research

Using different time frame and several different projects, the study can provide very different results. Future studies could contain data from longer period and get even more data of competitor sales. The data set could be specified to contain only a certain type of apartments from competitors' projects. This would require massive manual work, but can be possible in the future. If the exact sales transactions of competitors are not available, even monthly data of supply and competition would bring some additional accuracy to the model. Furthermore, it would be interesting to study the effects on competition in different market situations, such as when interest rates are low or high. If taking into account more projects in the same area,

it would provide more comprehensive data, more observations and additional features explaining the prices and TOM, such as project specific factors and location. In addition, there could be plenty of other factors that may affect selling times, for example, as Kang and Gardner (1989) suggest the expected interest rates.

One perspective for future research could be finding an optimal pricing strategy while mirroring these factors that can affect sales. Further studies could be performed using a measure of mispricing, such as the degree of overpricing. By using the degree of overpricing instead of price in TOM model, the endogeneity problem can be tackled. This could bring more information on the relationship of price and TOM to get more clear vision what could be the optimal pricing strategy. The pricing strategy model can also consider the efforts for marketing and brand building.

REFERENCES

Anglin, P., Rutherford, R. & Springer, T. (2003) The trade-off between the selling price of residential properties and time-on-the-market: The impact of price setting. The Journal of Real Estate Finance and Economics 26(1), 95-111

Asabere, P., Huffman, F. & Mehdian, S. (1993). Mispricing and Optimal Pricing on the Market. The Journal of Real Estate Research, 8(1), 149-156

Asabere, P. & Huffman, F. (1992) Macro-Determinants of Time on the Market. Real Estate Issues. 17(1), 39-43

Atterhög, M., & Lind, H. (2004) How does increased competition on the housing market affect rents? An empirical study concerning Sweden. Housing studies, 19(1), 107-123

Baranzini, A. & Ramirez, J. (2005) Paying for quietness: The impact of noise on Geneva rents. Urban Studies, 42(4), 633–646

Barlindhaus, R. & Nordahl, B. (2017) Developers' price setting behaviour in urban residential redevelopment projects. Journal of European Real Estate Research, 11(1)

Bartik, T. & Smith, V. (1987) Urban amenities and public policy. In Handbook of regional and urban economics,2, 1207-1254

Belkin, J., Hempel, W. & McLeavey, D. (1976) An empirical Study of Time on the Market Using Multidimensional Segmentation of Housing Markets. American Real Estate and Urban Economics Association Journal, 4(2), 57-75

Blatná, D. (2006). Outliers in regression. Trutnov, [Online article]. [Referred 20.3.2022] Available https://statistika.vse.cz/konference/amse/PDF/Blatna.pdf

Bratislava II (2021) Wikipedia [Online article]. [Referred 29.12.2021] available: https://en.wikipedia.org/wiki/Bratislava_2

Bollen, K. & Jackman, R. (1990). Regression Diagnostics: An Expository Treatment of Outliers and Influential Cases. Modern Methods of Data Analysis, 13(4), 257-291

Bulan, L., Mayer, C. and Somerville, T. (2009) Irreversible Investment, Real Options, and Competition: Evidence from Real Estate Development. Journal of urban economics 65(3), 237–251.

Can, A. (1992) Specification and estimation of hedonic housing price models, Regional Science and Urban Economics, 22(3), 453-74

Capozza, D., Hendershott, P., Mack, C. & Mayer, C. (2002) Determinants of real house price dynamics. NBER Working Paper, No. W9262

Cár, M. (2006) Monitoring residential property prices in Slovakia. Biatec 14(1), 9-15

Casson, R. & Lachlan, F. (2014) Understanding and Checking the Assumptions of Linear Regression: a Primer for Medical Researchers: Assumptions of Linear Regression. Clinical & experimental ophthalmology, 42(6), 590–596

Cesnak, M. & Klacso, J. (2021) Assessing real estate prices in Slovakia - a structural approach, 3, Research Department, National Bank of Slovakia.

Chang, C. & Ward, C. (1993). Forward pricing and the housing market: The pre-sales housing system in Taiwan. Journal of Property Research, 10(3), 217-227

Chin, T. & Chau, K. (2003) A critical review of literature on the hedonic price model, International Journal for Housing and Its Applications 27(2), 145-165

Chu, Y. & Sing, T. (2007) Optimal timing of real estate investment under an asymmetric duopoly. Journal of Real Estate Finance and Economics, 34(3), 327–345.

Ciarlone, A. (2015) House price cycles in emerging economies. Studies in Economics and Finance, 32(1), 17–52

Crowd Estate (2021) Development of real estate prices in Bratislava and its surroundings. [Online article]. [Referred 8.2.2022]. Available https://blog.crowdestate.eu/en/development-of-real-estate-prices-in-bratislava-and-its-surroundings/

Cubbin, J. (1974) Price, quality and selling time in the housing market, Applied Economics, 6(3), 171–187

Datacube (2021a) DATAcube Statistics: Indices of Realized Prices of Dwellings – quarterly. [Online source]. [Referred 3.1.2021]. Available http://datacube.statistics.sk/#!/view/en/VBD_SK_WIN/sp1002qs/v_sp1002qs_00_00_00_en

Datacube (2021b) DATAcude Statistics: Housing construction - selected indicators [Online source]. [Referred 5.1.2021]. Available http://datacube.statistics.sk/#!/view/en/VBD_SK_WIN/st3004rr/v_st3004rr_00_00_00_en

DiPasquale, D. (1999). Why don't we know more about housing supply?. The Journal of Real Estate Finance and Economics, 18(1), 9-23

DiPasquale, D. & Wheaton, W. (1994) Housing Market Dynamics and the Future of Housing Prices. Journal of Urban Economics, 35(1), 1-27

Dluhoš, M. (2017) An econometric model of real estate prices in Slovakia. Faculty of Business and Economics, Brno, 30(November), 44-52

Dormann, C., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J., Gruber, B., Lafourcade, B., Leitão, P. & Münkemüller, T. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance, Ecography, 36(1), 27-46

Dubé, J. & Legros, D. (2016). A Spatiotemporal Solution for the Simultaneous Sale Price and Time-on-the-Market Problem. Real Estate Economics, 44(4), 846-877

(EHS) Engineering Statistics Handbook (2022) Exploratory data analysis: Measures of skewness and Kurtosis. [Online book]. [Referred 19.2.2022]. Available https://www.itl.nist.gov/div898/handbook/eda/section3/eda35b.htm#:~:text=The%20skewness%20for%20a%20normal,data%20that%20are%20skewed%20right.

European Commission (2021) European Construction Sector Observatory: Country profile Slovakia. September. [Online article]. [Referred 5.1.2021]. Available https://ec.europa.eu/growth/system/files/2021-11/ECSO_CFS_Slovakia_2021.pdf

Fletcher, M. Gallimore, P. and Mangan, J. (2000) Heteroscedasticity in hedonic house price models, Journal of Property Research, 12(2), 93-108

Follain, J. (1979) The price elasticity of the long-run supply of new housing construction. Land Economics, 55(2), 190-199

Forrest, J. (1996) The impact of a light rail system on the structure of house prices. Journal of Transport Economics and Policy, 31(4), 15-29

Frost, J. (2021) How to Interpret P-values and Coefficients in Regression Analysis. Statistics by Jim. [Online article]. [Referred 29.12.2021]. Available https://statisticsbyjim.com/regression/interpret-coefficients-p-values-regression/#:~:text=The%20sign%20of%20a%20regression,variable%20also%20tends%20to%20increase.

Gallin, J. (2006) The long-run relationship between house prices and income: Evidence from local housing markets. Real Estate Economics, 34 (3), 417–438

Garrod, K. (1992) Valuing the goods characteristics – an application of the hedonic price method to environmental attributes. Journal of Environmental Management, 34 (1), 59-76

Gustavsson, F. & Vahtola, S. (2014) Pricing Strategies in newly developed housing projects. KTH Architecture and the Built Environment, Department of Real Estate and Construction Management. Stockholm.

Golej, J., Pánik, M. & Špirková, D. (2016) Post-crisis development of residential real estate prices and their determinants of influence in Slovakia. International Multidisciplinary Scientific GeoConference: SGEM, 3, 441-451

Goodman, A. (1998) Andrew Court and the Invention of Hedonic Price Analysis. Journal of Urban Economics. 44(2): 291-298. Available on Internet: https://doi.org/10.1006/juec.1997.2071

Greenstone, M. (2017) The continuing impact of Sherwin Rosen's "Hedonic prices and implicit markets: product differentiation in pure competition". Journal of Political Economy, 125(6), 1891-1902

Grenadier, S. (2002) Option exercise games: an application to the equilibrium investment strategies of firms. Review of Financial Studies, 15(3), 691–721

Halvorsen, R. & Pollakowski, H. (1981) Choice of Functional Form for Hedonic Price Equations. Journal of Urban Economics, 10(1), 37-49

Haurin, D. (1988) The duration of marketing time of residential housing, Real Estate Economics, 16(4), 369-410

Healey, P. (1994) Urban policy and property development: the institutional relations of real estate development in an old industrial region. Environment and Planning A, 26(2), 177-198.

Holm. E. (2021) Introduction to Research Methods. [Online book]. [Referred 28.2.2022] Available on https://bookdown.org/ejvanholm/Textbook/

Ho, D., Hui, E., Ho, T. & Rengarajan, S. (2019) Strategic behavioral pricing of the private residential development market–a simplified experimental approach. International Journal of Housing Markets and Analysis.12(4), 526-557

Hort, K. (1998) The determinants of urban house price fluctuations in Sweden 1968–1994. Journal of Housing Economics, 7(2), 93–120

Horvatova, E. (2020) Twenty Years of Mortgage Banking in Slovakia. International Journal of Financial Studies, 8(3), 56.

Hua, C. Chang C. & Hsieh, C. (2001) The price volume relationships between the existing and the pre-sales housing markets in Taiwan. International Real Estate Review. 4(1), 80-94

Hui, E., Lian, C., Wang, Z. & Wang, Y. (2016) The roles of Developers's status and competitive intensity in presale pricing in a residential market: A study of the spatio-temporal model in Hangzhou, China. Urban Studies Journal, 53(6), 1203-1224

Hui, E. & Hung Yu, K. (2012) The impact of pricing strategies on time-on-market under various economic conditions. International Journal of strategic property management, 16(1), 56-70

Hwang, M. & Quigley, J. (2006) Economic fundamentals in local housing markets: evidence from U.S. metropolitan regions, Journal of Regional Science, 46(3), 425–453

Jud, G., Winkler, D. & Kissling, G. (1995) Price spreads and residential housing market liquidity, Journal of Real Estate Finance and Economics, 11(3), 251–260

Kalra, R. & Chan, K., (1994). Censored sample bias, macroeconomic factors, and time on market of residential housing. Journal of Real Estate Research, 9(2), 253– 262

Kang, H.B., Gardner, M.J., 1989. Selling price and marketing time in the residential real estate market. Journal of Real Estate Research, 4(1), 21-35

Knight, J. (2002) Listing Price Time on Market, and Ultimate Selling Price: Causes and Effects of Listing Price changes. Real Estate Economics. 30(2), 213-37

Kuryj-Wysocka, O & Radosław, W. (2013) The problem of competition on real estate market on the example of the local housing market in Olsztyn. Real Estate Management and Valuation, 21(1), 69-76

Laakso, S. (1997) Urban Housing Prices and the Demand for Housing Characteristics. A Study on Housing Prices and the Willingness to pay for Housing Characteristics and Local Public Goods in the Helsinki Metropolitan Ar, ETLA A, The Research Institute of the Finnish Economy, 27(December)

Lai, R, Ko, W. & Yuqing, Z. (2004) Sale before completion of development: pricing and strategy. Real Estate Economics, 32(2), 329-357

Lancaster, K. (1966) A new approach to consumer theory. Journal of political economy, 74(2), 132-157

Łaszek, J., Olszewski, K., & Waszczuk, J. (2016) Monopolistic Competition, Price Discrimination as a Development Company Behaviors in the Housing Primary Market. In The Narodowy Bank Polski Workshop: Recent trends in the real estate market and its analysis.

Leszczyński, R., & Olszewski, K. (2017). An analysis of the primary and secondary housing market in Poland: evidence from the 17 largest cities. Baltic Journal of Economics, 17(2), 136-151.

Li, W (2004) The impact of pricing on time-on-the-market in high-rise multiple-unit residential developments. Pacific Rim Property Research Journal, 10(3), 305-327

Lexxus (2020a) Lexxus Research: Residential Real Estate Market Q1 2020. Bratislava

Lexxus (2020b) Lexxus Research: Residential Real Estate Market Q4 2020. Bratislava

Lexxus (2016-2021) Lexxus Research: Residential Real Estate Market Q4 2020. Bratislava

Liptáková, J. (2021) Bratislava property prices soar amid pandemic. Slovak Spectator. [Online article]. [Referred 4.1.2022]. Available https://spectator.sme.sk/c/22667117/bratislava-property-prices-soar-amid-pandemic.html

Liptáková, J. (2018) Rental housing has become a hot issue. Slovak Spectator. [Online article]. [Referred 4.1.2022]. Available https://spectator.sme.sk/c/22419944/rental-housing-has-become-a-hot-issue.html

Ma, L., Liu, H. & Sing, M. (2020) Responsiveness of residential construction-production progress to house price dynamics. International Journal of Housing Markets and Analysis. 14(4), 796-814

Malpezzi, S. (2003) Hedonic pricing models: a selective and applied review. Housing economics and public policy, 1, 67-89

Mayer, C. (1995) A Model of Negotiated Sales Applied to Real Estate Acutions. Journal of Urban Economics 38(1), 1-22

Miller, N. (1978) Time on market and selling prices, Real Estate Economics, 6(2), 164-174

Mikhed V & Zemčík, P. (2007). Do house prices reflect fundamentals? Aggregate and panel data evidence. Journal of Housing Economics, 18(2), 140-149

Mok, D. (2003) Disentangling the Market Price: Seller's Reservation Price, Buyer's Offer Price, and the Listing's Days on the Market. Working paper, McGill University

Montgomery, D., Peck, E. & Vining, G. (2012). Introduction to linear regression analysis John Wiley & Sons.

NBS (2022) Macroeconomic database: Financial Market: Interest rates to loans for Households. [Online database]. [Referred 8.2.2022]. Available https://www.nbs.sk/en/monetary-policy/macroeconomic-database/macroeconomic-database-chart

O'Brien, R. (2007) A Caution Regarding Rules of Thumb for Variance Inflation Factors. Quality & Quantity, 41, 673–690

Ong, S. & Koh, Y. (2000) Time-on-the-market and price trade-offs in high rise housing sub-markets, Urban Studeies, 37(11), 2057-2071

Ooi, J. & Thao, T. (2012) New supply and price dynamics in the Singapore housing market. Urban Studies 49(7), 1435-1451

Ott, S., Hughen, W., & Read, D. (2012) Optimal phasing and inventory decisions for large-scale residential development projects. The Journal of Real Estate Finance and Economics, 45(4), 888-918

Owusu-Ansah, A. (2011) A review of Hedonic Pricing Models in Housing Research. Journal of international real estate and construction studies 1(1), 19-39

Pakarinen, S. (2018) Semiparametric Efficiency Analysis of Housing Markets. Doctoral thesis. Aalto University School of Business, Finland. [Online article]. [Referred 12.2.2022]. Available https://aaltodoc.aalto.fi/handle/123456789/30197

Porter, M. (1985) Competitive Advantage: Creating and Sustaining Superior Performance. New York, NY: Free Press.

Porter, M. (1980) Competitive Strategy: Techniques for analyzing industries and competitors. Free Press, New York/ Collier Macmillan, London.

Quigley, J. (1979) What Have We Learned About Urban Housing Markets? Current Issues in Urban Economics Baltimore: Johns Hopkins University Press, 13(3), 391-429

Rahmatian, M. & Cockerill, L. (2004) Airport noise and residential housing valuation in Southern California: a hedonic pricing approach, International Journal of Environmental Science and Technology, 1(1), 17-25

Ramsey, J. (1969) Tests for specification errors in classical linear least-squares regression analysis. Journal of the Royal Statistical Society Series B, 31(2), 350–371

Rehák, Š., & Káčer, M. (2019). Estimating price gradient in Bratislava with different distance measurements. Journal of European Real Estate Research.

Rosen, S. (1974) Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. Journal of Political Economy. 82(1), 34-55

Rybárová, D., Braunová, M. & Jantošová, L. (2016). "Analysis of the construction industry in the Slovak Republic by bankruptcy model." Procedia-Social and Behavioral Sciences, 230, 298-306.

Schwartz, E. & Torous, W. (2003) Commercial Office Space: Tests of a Real Options Model with Competitive Interactions. Working Paper. University of California–Los Angeles

Simons, R., Quercia, R. & Maric, I. (1998) The value impact of new residential construction and neighborhood disinvestment on Residential Sale price. Journal of Real Estate Research. 15(1), 147

Singla, H. & Priyanka, B. (2019) Factors affecting rentals of residential apartments in pune, india: An empirical investigation. International Journal of Housing Markets and Analysis, 12(6), 1028-1054

Sirmans, S., MacDonald, L. & Macpherson, D. (2010) A meta-analysis of selling price and time-on-the-market. Journal of Housing Research, 19(2), 139-152

Sirmans, S., Macpherson, D., & Zietz, E. (2005). The composition of hedonic pricing models. Journal of real estate literature, 13(1), 1-44

Sirmans. C., Tumbull, G. & Dombrow, J. (1997) Residential Development. Risk, and Land Prices. Journal of Regional Science, 37(4), 613-628

Somerville, C. (2001) Permits, Starts, and Completions: Structural Relationships versus Real Options. Real Estate Economics 29(1), 161–190

Sozewica (2021) Box-Cox transformaitons is the magic we need. Towards data science. [Online article]. [Referred 1.3.2022]. Available https://towardsdatascience.com/box-cox-transformation-is-the-magic-we-need-3d0ef9a18339

Spirkova, D. (2018). Housing Policy in the Slovak Republic. In Housing. IntechOpen.

Springer, T. (1996) Single-Family Housing Transactions: Seller Motivations, Price and Marketing Time. Journal of Real Estate Finance and Economics, 13(2), 237–54

Stevenson, S. (2004) New empirical evidence on heteroscedasticity in hedonic housing models, Journal of Housing Economics, 13(2), 136-153

Studenmund, A. (2010) Using econometrics : a practical guide, 6th ed (9780131367739). Prentice Hal

Swedish Competition Authority (2002a) De Svenska priserna kan pressas Report, 5, Stockholm

Swedish Competition Authority (2002b) Konkurrensen I Sverige Report, 4, Stockholm

Taylor, C. (1999) Time-on-the-market as a sign of quality. Review of Economic Studies, 66(3), 555-578

Thanasi, M. (2015) Hedonic appraisal of apartments in Tirana. International Journal of Housing Markets and Analysis, 9(2), 239-255

Tobler W. (1970) A computer movie simulating urban growth in the Detroit region. Economic Geography, 46, 234–240

Topel, R. & Rosen, S. (1988) Housing investment in the United States. Journal of Political Economy. 96(4), 718-740

UC (2022) UC Business Analytics R Programming Guide. Github. [Online article]. [Referred 20.3.2022]. Available https://uc-r.github.io/assumptions_normality#shapiro

United Nations (1999) Country profiles on the housing sector: Slovakia. Economic commission for Europe. Sales No. E. 99.11.E.20. Geneva [Online article]. [Referred 4.1.2022]. Available https://unece.org/DAM/hlm/prgm/cph/countries/slovakia/CP_Slovakia_en.pdf

Volkova, V. & Pankina, V. (2013) The research of distribution of the Ramsey RESET-test statistic. International Conference on Actual Problems of Electronics Intstrument Engineering.

Wang, Y., Tang, W., & Jia, S. (2016). Uncertainty, competition and timing of land development: Theory and empirical evidence from Hangzhou, China. The Journal of Real Estate Finance and Economics, 53(2), 218-245

Wheaton, W. (1990) Vacancy, Search and Prices in a Housing Market Matching Model. Journal of Political Economy, 98, 1270-1293

Weimer, A. & Hoyt, H. (1948) Principles of Urban Real Estate. Ronald Press Co., New York.

Williams, J. (1993) Equilibrium and Options on Real Assets, Review of Financial Studies, 6(4), 852-850

Wong, S., Chau, K., Yau, Y. & Cheung, A. (2011) Property price gradients: the vertical dimension, Journal of Housing and the Build Environment, 26(1), 33-45

Yang, J. (2020) Interpreting Regression Coefficients for Log-Transformed Variables. Cornell Statistical Consulting Unit. [Online article]. [Referred 28.2.2020]. Available https://cscu.cornell.edu/wp-content/uploads/83_logv.pdf

Yang, S & Yavas, A. (1995a) Bigger is not better: brokage and time-on-the-market, Journal of Real Estate Research, 10(1), 23-33

Yavas, A. and Yang, S. (1995b). The strategic role of listing price in marketing real estate: theory and evidence. AREUEA Journal, 23(3), 347–368

Yavas, A. (1992) A Simple Search and Bargaining Model of Real Estate Markets. Journal of the American Real Estate and Urban Economics Association, 20(4), 553-48

Yu, H. (2010) China's house price: Affected by economic fundamentals or real estate policy? Frontiers of Economics in China; Beijing, 5(1), 25-51

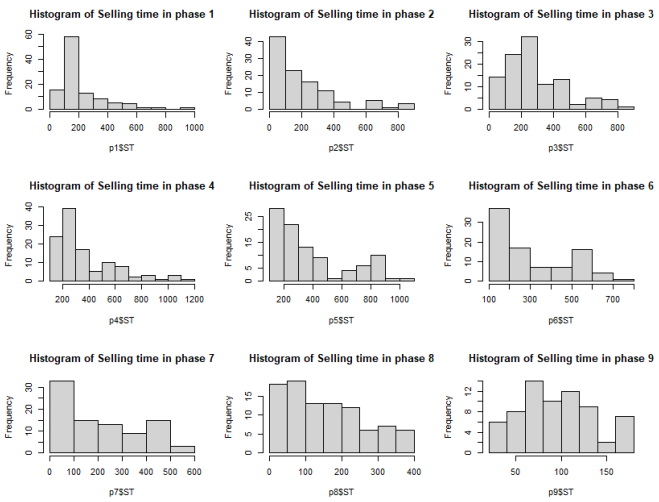
Zhang, X. (2010) Competitiveness for real estate developers: a China study. Department of Building and Real Estate. Hong Kong Polytechnic University.

Zhou, X., Zahirovic-Herbert, V., & Gibler, K. (2018) Time-on-market in Chinese condominium presales. International Journal of Strategic Property Management, 22(3), 191-203

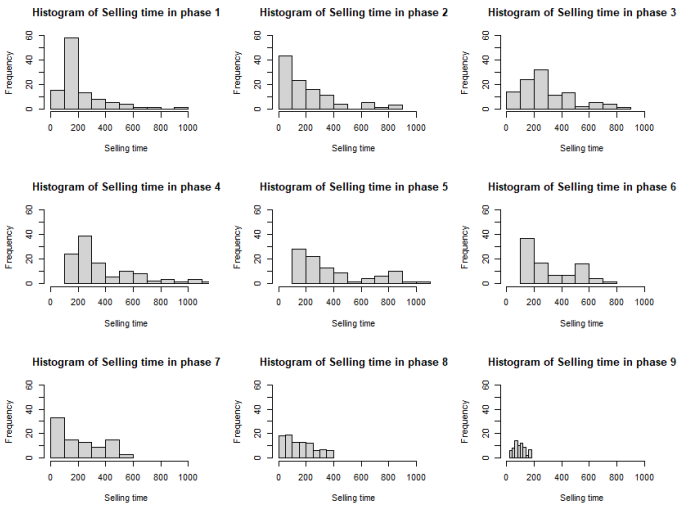
APPENDICES

Appendix 1: Histograms of TOM in each phase

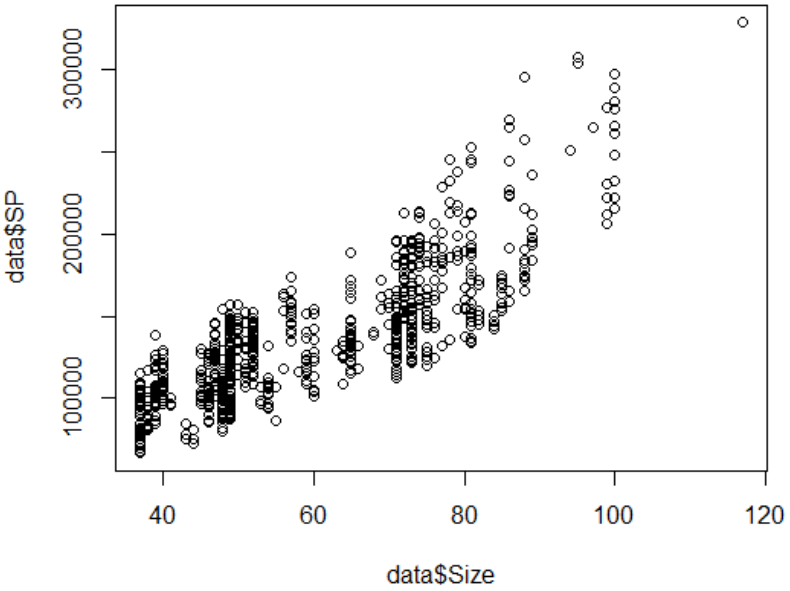
(a) Histogram of TOM in each phase



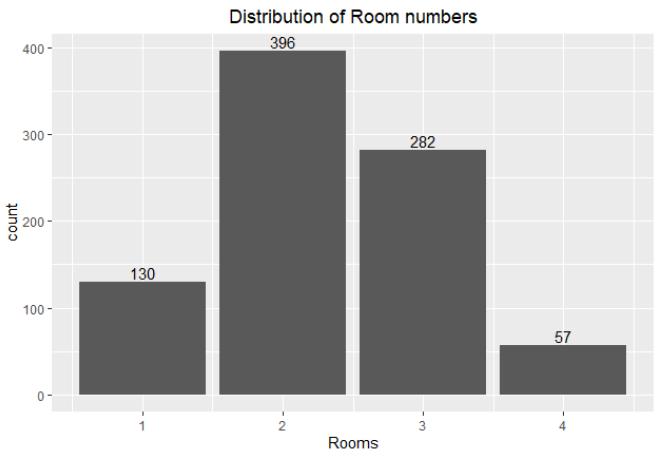
(b) Histograms of TOM in each phase, scaled to same axis



Appendix 2: Scatter plot of selling price and size



Appendix 3: Distribution of rooms

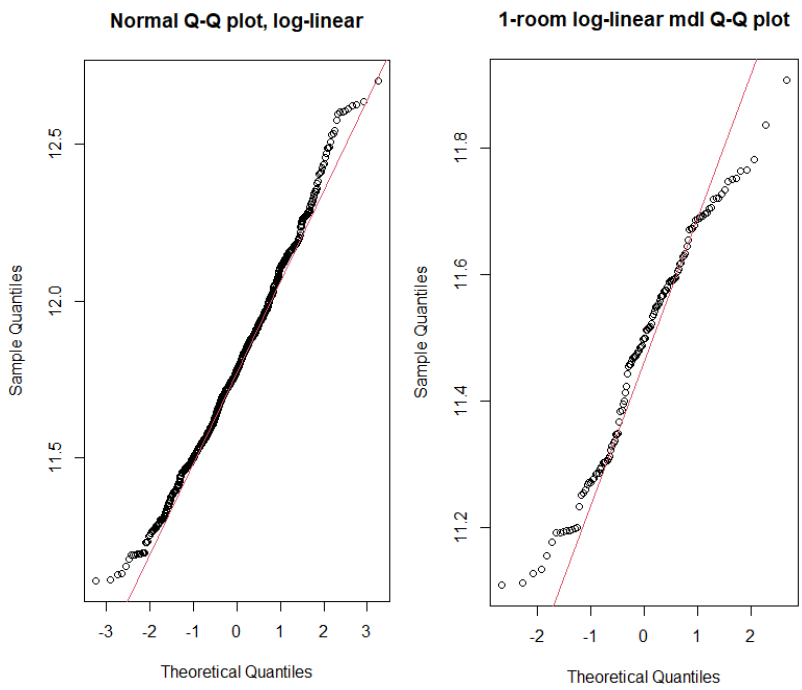


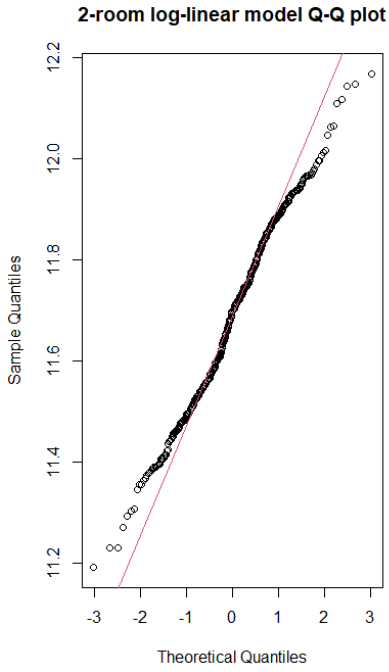
Appendix 4: Results for additional two-room model, without competition dummy variables and with season variables

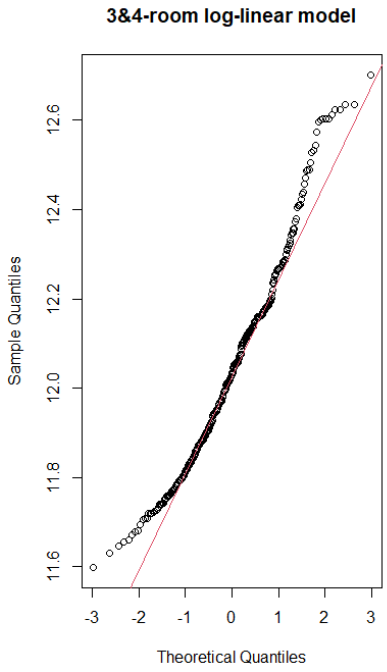
	Estimate		Std. Error	t value	Pr(>|t|)
(Intercept)	7,85724	***	0,5914	13,287	< 2e-16
Size	-5,0804	***	0,01058	-4,929	1,23E-06
Ground	2,266299		0,0801	0,28	0,77977
High	35,67607	**	0,08897	3,43	0,00067
Top	-38,9453		0,4755	-1,038	0,3
Balcony_size	-5,36852	**	0,01734	-3,183	0,00158
Terrace_size	-1,85557	**	0,006958	-2,692	0,00741
SP	0,003329	***	2,86E-06	11,633	< 2e-16
Summer	42,80359	***	0,08503	4,19	3,47E-05
Fall	60,36784	***	0,08342	5,661	2,95E-08
Cprice_type	-0,00269	***	2,44E-06	-11,058	< 2e-16
Supply_type	-0,12932	**	0,000437	-2,965	0,00322

Appendix 5: Q-Q plot of standardized residuals

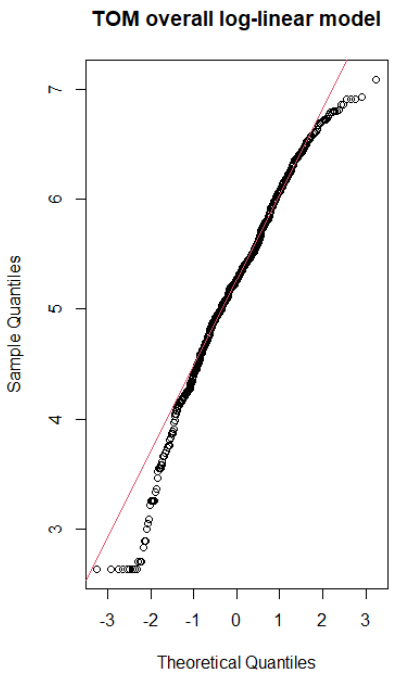
(a) Price models

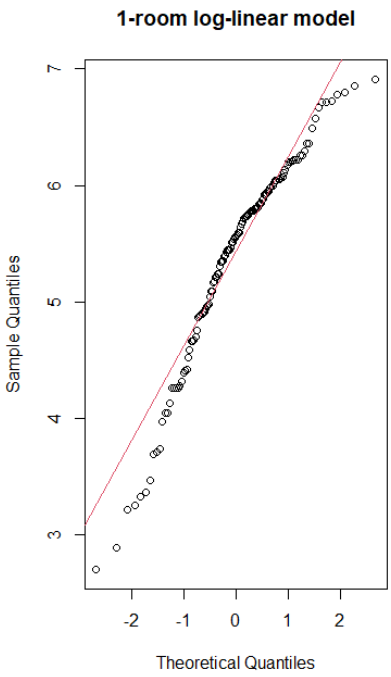


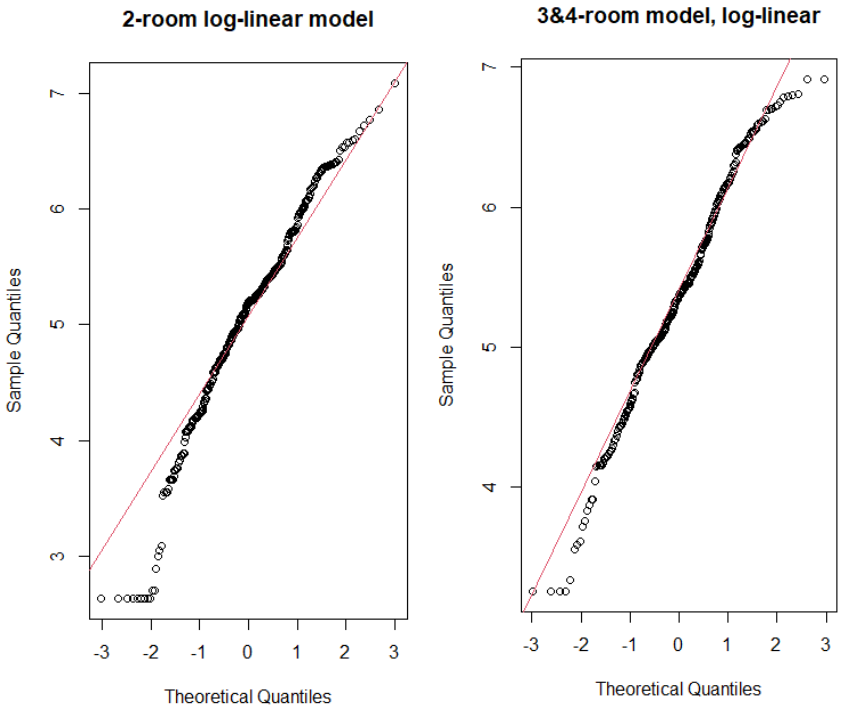

2-room log-linear model Q-Q plot
Sample Quantiles
Theoretical Quantiles
11.2
11.4
11.6
11.8
12.0
12.2
-3
-2
-1
0
1
2
3


3&4-room log-linear model
Sample Quantiles
Theoretical Quantiles
11.6
11.8
12.0
12.2
12.4
12.6
-3
-2
-1
0
1
2
3

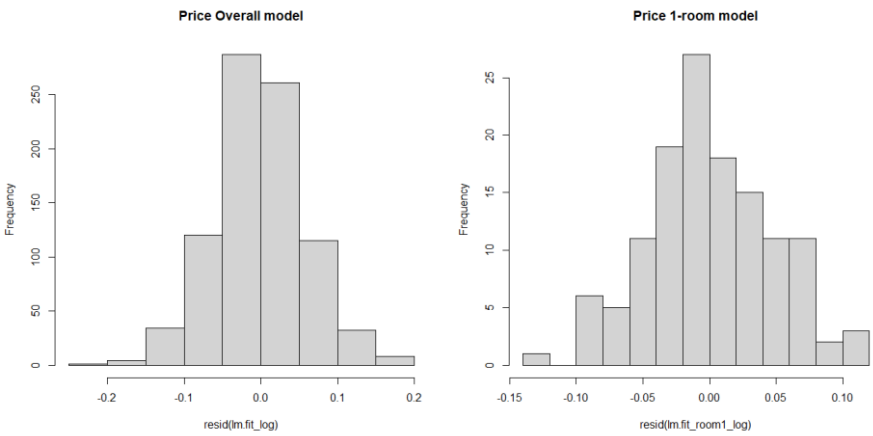
(b) TOM Models

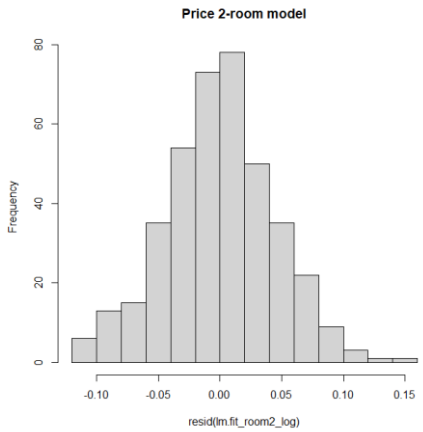

TOM overall log-linear model
Sample Quantiles
Theoretical Quantiles
3
4
5
6
7
-3
-2
-1
0
1
2
3

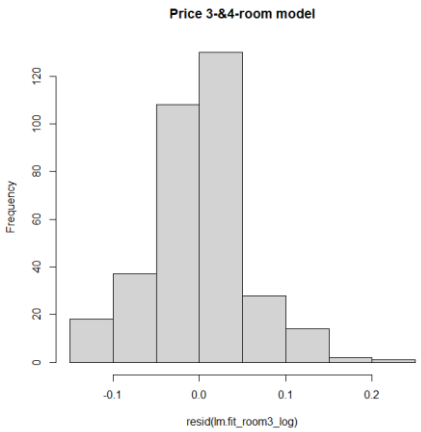

1-room log-linear model
Sample Quantiles
Theoretical Quantiles
3
4
5
6
7
-2
-1
0
1
2

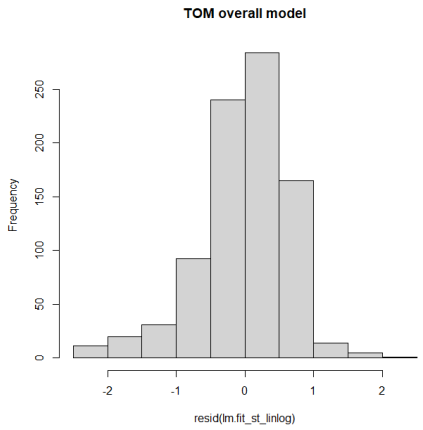


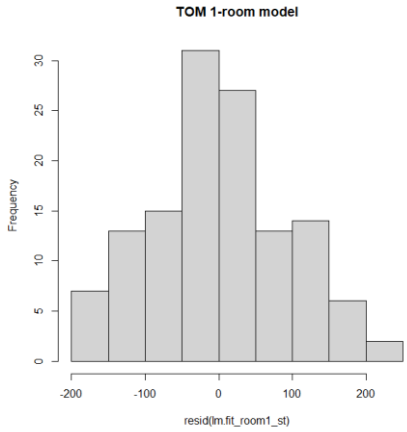
Appendix 6: Histogram of residuals

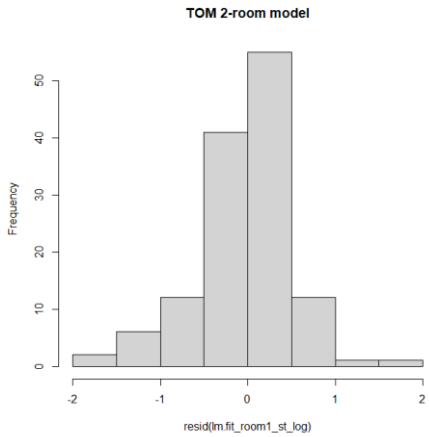


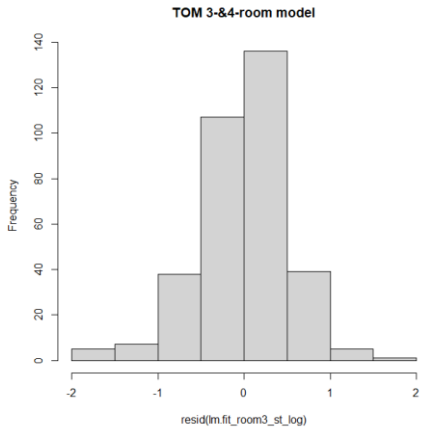

Price 2-room model
Frequency
0
20
40
60
80
-0.10
-0.05
0.00
0.05
0.10
0.15
resid(lm.fit_room2_log)


Price 3-&4-room model
Frequency
0
20
40
60
80
100
120
-0.1
0.0
0.1
0.2
resid(lm.fit_room3_log)


TOM overall model
Frequency
0
50
100
150
200
250
-2
-1
0
1
2
resid(lm.fit_st_linlog)


TOM 1-room model
Frequency
0
5
10
15
20
25
30
-200
-100
0
100
200
resid(lm.fit_room1_st)


TOM 2-room model
Frequency
0
10
20
30
40
50
-2
-1
0
1
2
resid(lm.fit_room1_st_log)


TOM 3-&4-room model
Frequency
0
20
40
60
80
100
120
140
-2
-1
0
1
2
resid(lm.fit_room3_st_log)

Appendix 7: Comparison of linear, log-linear and box-cox transformed price models

	Linear overall price model	Log-linear overall price model	Boc-cox overall price model
R^2	0.9492	0.9595	0.9627
Ramsey RESET test	0.0000	0.0000	0.0000

	1-room models			2-room models			3-&4-room models		
	Linear	Log-linear	Box-cox	Linear	Log-linear	Box-cox	Linear	Log-linear	Box-cox
R^2	0.9200	0.9354	0.9414	0.9372	0.9414	0.9415	0.9192	0.9373	0.9379
Ramsey's RESET test	0.0001	0.0435	0.1612	0.0014	0.697	0.198	0.0000	0.4309	0.000967

Appendix 8. Comparison of linear, log-linear and log-log TOM models

	TOM overall linear model	TOM overall log-linear model	TOM overall log-log model
R^2	0.4267	0.3800	0.4001
Ramsey RESET test	0.0000	0.0001	0.08853

	1-room models			2-room models			3-&4-room models		
	Linear	Log-linear	Log-log	Linear	Log-linear	Log-log	Linear	Log-linear	Log-log
R^2	0.7860	0.6902	0.7037	0.7157	0.5495	0.5744	0.5384	0.5289	0.5799
Ramsey's RESET test	0.0001	0.0003	0.0005	0.0000	0.18	0.3196	0.0000	0.1484	0.0000