Institute of Economic Studies
Faculty of Social Sciences
Charles University

REGIONAL DETERMINANTS OF HOUSING PRICES IN THE CZECH REPUBLIC

Roman Kalabiska
Michal Hlavacek

IES Working Paper 9/2020

Institute of Economic Studies,
Faculty of Social Sciences,
Charles University in Prague

[UK FSV – IES]

Opletalova 26
CZ-110 00, Prague
E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Institut ekonomických studií
Fakulta sociálních věd
Univerzita Karlova v Praze

Opletalova 26
110 00 Praha 1

E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Bibliographic information:
Kalabiska R. and Hlavacek M. (2020): "Regional Determinants of Housing Prices in the Czech Republic" IES Working Papers 9/2020. IES FSV. Charles University.

This paper can be downloaded at: http://ies.fsv.cuni.cz

Regional Determinants of Housing Prices in the Czech Republic

Roman Kalabiska[a]
Michal Hlavacek[a,b]

[a] Institute of Economic Studies, Faculty of Social Sciences, Charles University
Opletalova 26, 110 00, Prague, Czech Republic
[b] Office of the Czech Fiscal Council, Holeckova 103/31, 15000 Prague 5, Czech Republic
Email (corresponding author): r.kalabiska@seznam.cz

April 2020

Abstract:
This paper examines the behaviour of housing prices and identifies their determinants across Czech regions from 2000 to 2017. The effect of a wide range of variables on apartment prices is analyzed on quarterly data for all regions of the Czech Republic using panel dynamic OLS estimator. Furthermore, an error correction model is employed to verify the existence of long-term equilibrium of apartment prices and the speed of price adjustment in the short run. The regression reveales that apartment prices are driven mainly by wages, unemployment rate and building plot prices. In order to check robustness of selected model, several regions with unique characteristics are excluded from the sample and analyzed separately. Our results show that building plot prices have an unexpected negative effect in low-income regions and labour force factors are less important in Prague, caused by a number of unique features of the capital city.

JEL: C23, O18, R11, R13, R31
Keywords: apartment prices, regional analysis, residential real estate, panel regression

1 Introduction

The ownership of residential property is one of the key components of household wealth. It offers an opportunity to accumulate assets and build wealth and thus through wealth effect influences household consumption and investment decisions. When buying a property, one has to focus not only on its price, but also its characteristics. Each property has unique features which define its value. One of the most important factors that affect property price is location. That is why in this paper, special attention is paid to the region in which a property is located. In each of the 14 regions of the Czech Republic we will define the determinants of residential real estate prices and estimate to what extent they affect these prices. Moreover, we will perform several robustness checks in order to show which role various factors play in different regions and unlike previous studies of housing price determinants, how relationships between them change over time.

The purchase of a property is a major component of household expenditures, therefore understanding the dynamics of real estate prices and its determinants is crucial for residents. However, because residential property often requires external financing by a mortgage as well as householders' own funds, and constitutes a type of collateral for private credit, the relationship between the real estate market and the financial sector is particularly important. In addition, assets whose value is linked to residential real estate are an important component of portfolios of financial intermediaries. The behaviour of property prices influences their profitability and the performance of the financial sector. Yusupova (2016) emphasizes that on a large scale, residential property financing and sharp price corrections can undermine financial stability and lead to a slowdown in the economic activity.

These interconnections became apparent after the financial crisis of 2007/2008. Due to low interest rates and a complex securitization of subprime mortgage loans, the subsequent high default rate of these loans, especially in the United States, led to the burst of the so-called "subprime bubble", which fueled an unprecedented growth of real estate prices. This bubble is usually mentioned among the most important factors of this crisis (Hlaváček and Komárek 2009b). An important question is whether an unusual rise in property price can be explained by fundamental factors or other variables. When the price is driven by speculation or some irrational factors, the creation of a bubble is a real possibility. Considering the devastating effects of the recent property price boom and bust on the world economy, monitoring the dynamics of real estate prices and the factors driving the price movements has become an important task for central banks in order to maintain financial stability (Tsatsaronis and Zhu 2004).

Against this background, this paper contributes to the research of real estate prices. A primary objective is to define factors which determine the price of residential real estate. As opposed to existing empirical literature, this paper also examines their dynamics over time and identifies events which changed these prices and overall functioning of real estate markets, such as the global financial crisis and consecutive low interest rate regime. The analysis uses a panel dataset covering a wide range of supply, demand, demographics and other macroeconomic variables across 14 Czech regions. Typically, studies of determinants of real estate prices collect data from different sources which can potentially result in an inconsistent dataset due to the differences in definitions of variables. Data in this paper are mostly collected from two sources, the Czech National Bank (CNB) and the Czech Statistical Office (CZSO), which potentially minimizes the bias caused by data heterogeneity. The results might be used for predicting property prices by central banks for testing various scenarios related to price stability. This paper should also shed some light on whether property prices respond more to the supply or demand side of the market, which is also important from the policy perspective.

This paper is organised as follows. The following chapter provides a brief overview of the real estate market in the Czech Republic, its participants and unique features. Chapter Three gives an overview of factors which play a role in determining real estate prices and discusses their expected effects. Chapter Four discusses empirical literature on the determinants and other studies of real estate prices, which may be helpful in the analysis conducted in this paper. Chapter Five discusses the problematics of data collection and provides the actual dataset and the theoretical framework for our analysis. Finally, the summary of results is presented and discussed in Chapter Six, along with a series of concluding remarks.

2 Real estate market in the Czech Republic

2.1 Market participants

As housing is a very complex commodity, there are many agents who enter the market. Ioannides and Rosenthal (1994) divided individuals who demand housing into two cathegories - **owner-occupiers** and **renters**. Tenants who pay rent to use the landlord's property, also called **renters**, do not create demand to buy real estate, they are pure consumers of it. The purchase of residential property is driven by both consumption and investment motives. One group of buyers are **households** who need to find housing to live in. Buying an owner-occupied housing is mainly driven by consumption motive, although it is inherently an investment into a fixed asset. Moreover, because households often offer their properties for rent, they can also act as **investors**. Therefore, these groups are not strictly separate.

Pure investors do not use the properties they own and usually wait for the most profitable investment. Sellers who need to sell their property quickly for any reason are often willing to lower the price, which creates a good opportunity for investors to make a profit, either from renting out the property or from re-selling it at a higher price. On the market of existing properties, supply is formed typically by **households** and **investors** who sell their properties because they want to move or to make a profit. New properties are supplied by **developers**. Although their main activities include buying land and building real estate, they can also renovate and re-lease existing properties. Due to the shortage of land and long construction times, the supply of new properties tends to be inelastic.

Other participants of real estate market include **real estate agencies**. They facilitate the rent, sale and purchase of a property, estimate its value, etc. Households usually do not possess enough money to finance a property by themselves and often take a mortgage loan from **banks**. As a result, both real estate agencies and banks indirectly affect the demand for housing. Furthermore, the **central bank** is responsible for implementing macro-prudential policy in order to reduce the build-up of risks in the financial sector. Last but not least, a significant role in the market is played by the **government**, which controls taxes, subsidises building savings and regulates the rights of tenants and landlords.

2.2 Market features

Real estate markets possess some unique characteristics, which influence the interaction between supply and demand and thereby the dynamics of property prices. These traits separate real estate from standard market commodities and make the study of real estate markets more complicated.

Lux (2009) emphasizes the **heterogeneity** of real estate. Unlike the common range of commodities that consumers can buy, each property is in terms of its characteristics unique. Individual houses and apartments differ in floor space, location (including the access to public infrastructure), or age. For two identically built apartments in one building, even the orientation within the building plays a role. Therefore, no two properties are fully alike, making them difficult to compare. Consequently, the **supply** of real estate is **inelastic**. In the secondary market, it is not possible to purchase a property if the current owner is not willing to sell it. As for new properties, Hilbers et al (2008) point out that the supply often reacts on changes in demand with a long lag due to the time needed to obtain construction permits, secure financing and finish construction. As a result, demand surplus leads to a persistent increase in prices, whereas a slowdown in demand leads to properties not being sold and remaining empty until the market forces adjust. As a result of this, the equilibrium is inherently not stable as the real estate market is not efficient in the short run.

Another unique characteristic of real estate is their **durability**. Buildings are constructed to last tens or hundreds of years and their useful life can be further extended by reconstruction. As a result, properties with significantly different ages can exist at the same time in one market. Based on the population census of 2011 (see CZSO, 2014), there were approximately 230 thousand houses (cca. 13% of existing stock) in the Czech Republic that were built before 1919, while almost 220 thousand were built in the period 2001-2011. Furthermore, the average age of houses was 49,8 years. For that reason, properties act as a long-term store of value and are therefore a popular type of investment.

As opposed to financial assets, housing is an **immobile good** characterized by **low liquidity** which cannot be moved from one location to another. To buy a property therefore means to buy the neighbourhood status, public infrastructure and a commute to the place of employment. As it is also a very expensive good, people tend to think carefully before they decide to make a purchase. Consequently, it usually takes longer to buy a property than to buy other consumer goods. (Lux 2009)

Next to the new construction, a significant part of real estate consists of the existing stock. However, due to **high transaction costs**, the number of transactions on the secondary market is relatively small. Since sales usually result from bilateral negotiations and third parties are involved, such as real estate agencies, banks or lawyers, there is a need for a **regulatory framework** to supervise the conditions of sales, mortgage loans or mortgage-backed securities. Due to the comprehensive structure of real estate market, the regulations are often quite complicated. Regulations go hand in hand with tax laws. **Taxation** can be set to favour either home ownership or renting, depending on the policy, and thus can strongly affect conditions in real estate markets. (Hilbers et al 2008)

3 Housing price determinants

In equilibrium, property price can be generally explained by fundamental factors, which affect both demand and supply. This chapter sets out to examine these and other factors, provides a discussion about the effects they are expected to have on real estate prices and explore potential problems that may be encountered in the analysis.

3.1 Supply factors

The supply of housing depends primarily on the profitability of construction firms. Following Égert and Mihaljek (2007), the overall cost of construction includes building plot prices, cost of material and wages of construction workers. Hlaváček and Komárek (2009b) divided the market of residential real estate into segments of existing and newly built properties. The supply of existing properties is inelastic and expected to have negative effect on property prices. We can use *the number of existing apartments* as a proxy for the **stock of existing properties**. In the segment of newly built properties, the price determines the amount of new construction. To measure **construction activity** in the segment of newly built properties, Belke and

Keil (2017) suggest to use *the number of newly constructed apartments*. However, the expected effect of construction activity is unclear. On one hand, construction increases supply relative to demand, suggesting a downward pressure on prices. On the other hand, new construction could be caused by increased demand for housing.

The relationship between prices of real estate and **building plot prices** is apparent. As Hlaváček and Komárek (2009b) explain in their paper, this variable suffers from an endogeneity problem. Each rise of building plot prices increases construction cost, which drives up the price of newly built apartments, but it also generates pressure on the price of old apartments. On the other hand, each rise of apartment prices incentivises the construction of new apartment houses, which fuels the demand for building plots and consequently their price increases as well. Therefore, to examine their relationship, it is necessary to use an endogenous model as the results of panel regression may be inaccurate. Both **cost of material** and **wages of construction workers** are expected to have a positive effect on property prices. As a proxy we will use *apartment construction prices*, which are provided by the Czech Statistical Office. They aggregate the total investment cost of building a new property and exclude the cost of the building plot. In their paper, Hlaváček and Komárek (2009b) established that the supply factors are not important in determining real estate prices. However, we also have to bear in mind that it takes a long time to construct buildings and go through the legislation process of obtaining a construction permit. That is why changes in supply factors affect real estate prices with a long lag.

3.2 Demand factors

Many researchers use **disposable income of households** as the main determinant of property prices. As increase in household income leads to the accumulation of wealth and increases the availability of a mortgage loan, a positive effect of income on property prices is expected. Above that, we expect stronger demand for housing and therefore higher prices in regions with higher income. In an econometric analysis, as a proxy variable can be used a macroeconomic indicator, such as *GDP per capita*, or a microeconomic indicator, usually *average monthly wage*. Since these variables tend to be highly correlated, Égert and Mihaljek (2007) suggest that it is favorable to choose only one of them.

Disposable income of households is directly affected by **labour market factors**, such as the unemployment rate, the economic activity of the population and the number of vacancies. Keeping the average wage constant, higher unemployment rate decreases the average disposable income, which leads to lower purchasing power of households and decrease in demand for housing. On the micro level, losing their jobs can force householders to substitute renting a property for buying one, thus decreasing demand for owner-occupied dwellings. The effect of economic activity rate and number of vacancies on property price is expected to be positive. Results of the study conducted by Hlaváček and Komárek (2009b) confirms the negative effect of *unemployment* on housing prices, while the *economic activity rate of the population* and the *number of vacancies* were insignificant. However, in Belke and Keil (2017), the unemployment rate turned out to be insignificant as well.

3.3 Demographic factors

Demographic factors describe the composition of population and affect housing prices either directly or indirectly through labour market. Furthermore, they also help to determine what types of properties are in demand. (Maher 2018)

An important demographic factor linked with aforementioned labour market is **population growth**. It can be either natural or caused by migration. The main motivator for migration is regional differences in real wages. An increase in net migration to a region inflates the demand for housing, which translates into higher prices. Hlaváček and Komárek (2009b) established the significance of both *natural population growth* and *population growth caused by migration*. Because of this, both variables will be used in this paper. They also found out that migration has stronger effect on property prices than household income, as not all immigrants join the workforce but all immigrants need housing. A positive relationship between population growth and real estate prices was confirmed by Dröes and van de Minne (2017), Capozza et al (2002), etc. Although a steady population growth *ceteris paribus* raises demand and thereby increases property prices, Li (2015) emphasizes that an excessive population growth can have a negative impact on the living environment and land capacity in large cities, which can eventually result in a decrease in property prices.

Belke and Keil (2017) also suggest that the *number of households*, the *population size* or the *size of labour force* in each region can be used instead. Due to low birth rate and high divorce rate in recent years, the average size of households has been decreasing and the number of households has been increasing faster than population. According to OECD (2011), this pattern has been observed in most developed countries. These variables, although correlated, would therefore produce different results.

Apart from the size of population, qualitative factors, such as age structure of the population or household formation play an important role in the real estate market. Productive age population forms the majority of demand for dwellings, as having a job is a precondition for owning a property. Regions with higher **share of productive age population** are expected to exhibit higher property prices due to stronger demand. Belke and Keil (2017) used the *dependency ratio* as a proxy for age structure defined as the number of persons aged under 15 or over 65 divided by the number of persons aged 15-65. In their study, the coefficient of age structure has the expected sign and is significant as a determinant of property prices. The same positive effect was also confirmed by Dröes and van de Minne (2017). In this paper we will use data for *the share of population aged 15-64* in each region published by CZSO.

Property prices are also expected to increase with a higher **divorce rate**, as most divorces turn one household into two, which creates a need for additional dwelling. This relationship was confirmed by Hlaváček and Komárek (2009b). The effect of **marriage rate,** however, is ambiguous, as a wedding can either establish a new household or merge two households into one. Moreover, the changes in the social attitude towards mariages may influence the data strongly. In the study conducted by Hlaváček and Komárek (2009b), marriage rate turned out to be statistically insignificant as a determinant of apartment prices.

3.4 Other factors

The relationship between **inflation** and property prices was established by Tsatsaronis and Zhu (2004), Apergis and Rezitis (2003), etc. The rise of commodity prices, particularly prices of building material and labour, increases the construction cost, which drives real estate prices up. However, Cohen and Karpavičiūtė (2017) used Granger causality test to show that the causality leads from housing prices to inflation, not *vice versa*.

A somewhat less common factor examined in academic literature is **international competitiveness**. An improvement of country's competitiveness due to lower inflation or a weak exchange rate relative to foreign countries can attract foreign investors and thereby increase domestic property prices. As a proxy Maher (2018) used *real effective exchange rate* (REER) and applied a vector autoregression (VAR) model to find that each percentage increase in REER results in 0,68% decline in real house prices in Sweden. He thereby concludes that gains in trade competitiveness lead to a significant appreciation of house prices. In czech literature, Hlaváček and Komárek (2009a) used the *ratio of foreign direct investments to GDP* to proxy demand from abroad, but observed only a weak effect. As for this paper, we will use the *real effective exchange rate* in the analysis. Another option would be the ratio of foreign dircet investment to GDP, which only offers yearly data. Both are published by the Czech National Bank and cover the whole country, not regions. The expected effect of REER on housing prices is negative, as increase in REER implies that exports became more expensive relative to imports. Thus, the country experiences a loss in trade competitiveness.

Interconnection with financial market

The real estate market is interconnected with the financial market through financial institutions. Demand for housing is largely determined by the accessibility of mortgage loans. A major factor that influences it is the **interest rate**. Cohen and

Karpavičiūtė (2017) argue that a higher interest rate increases the return of other assets, such as bonds, relative to the return of real estate, thus shifting demand from real estate to other fixed assets. However, the value of properties determines the value of collateral, which should be reflected in the average amount of mortgages. That is why interest rate might suffer from endogeneity. The effect of interest rate and number of mortgages on property prices has been established by Égert and Mihaljek (2007), Belke and Keil (2017), Maher (2018) and others. As for czech literature, Čadil (2009), Hlaváček and Komárek (2009b) used the number of mortgage loans in their studies, but did not confirm its significance as a determinant of real estate prices. Data on the *number of mortgage loans* are available at the Ministry of Regional Development (MMR). However, official data on mortgage rates published by CNB are available only since 2004. As a result, we will use a three-month rate of the interbank market (*3M PRIBOR*), which unfortunately does not include information on the spread of mortgage loans. This variable is cross-section invariant, i.e. the same for all regions.

Substitution between renting and home ownership

High mortgage rates and low rental prices can incentivise householders to live in a rented property rather than their own. Subsequent decline in demand for housing lowers their price. More importantly, if the rental price rises, the profitability of owning a property increases as well. This creates an incentive for speculators to purchase more properties whose price is then driven up. Therefore, a positive relationship between **rents** and property prices is expected. However, Hlaváček and Komárek (2009b) argue that real estate prices can also influence rents. The increase in prices on real estate market lowers the availability of owning a property while simultaneously increases demand for rental properties and thereby increases rents. Because of this, we expect rents to be endogenous.

Property features

One of many factors which affect residents' well-being is the **quality of available public infrastructure** in the area, in which their home is located. That includes the distance to the nearest public transport station, shopping centre, school for their children, etc. As a qualitative measure, this requires a suitable proxy variable. Belke and Keil (2017) used *the number of hospitals per 1,000 inhabitants* of each city analyzed in their research. They conclude that the public infrastructure is a significant determinant for apartment prices with a positive coefficient, which indicates higher demand in areas with higher quality infrastructure. They also suggest that *the number of secondary schools* can be used instead. In other available literature, Láznička (2016) analyzed the impact of *proximity to metro stations* on apartment prices in Prague and confirmed that with additional distance to the closest metro station the apartment price does generally decrease. In this paper, we will use the *number of hospital beds* in each region as a proxy for the quality of public infrastructure. Another measure, which may provide interesting results, is **urbanization**. The logic behind it is that public goods are more available in urban areas. Moreover, urban areas provide abundant job opportunities, which consequently manifests itself in the unemployment rate, wages, etc. To the authors' knowledge, the effect of urbanization on housing prices has not yet been tested in empirical literature.

Another qualitative factor which affects the market price of a property is its **technical condition**. That includes the age of the building, materials used on its construction, whether or not the property was reconstructed, etc. Other factors, as listed in Al-Marwani (2014) or Fujiwara and Campbell (2011), may include the **exposure to non-market goods**, such as pollution, noise or crime. However, due to the lack of official data, the use of this variable in economic literature is very limited.

Due to the complexity of real estate, there is a wide range of factors influencing their price. This chapter summarizes factors in terms of supply and demand, which are to be utilized in the empirical part of this paper. Additionally, commonly used factors are included to provide a comprehensive picture of the complexity of the real estate market. Although supply and demand for housing both interact to determine a long-term equilibrium price of real estate, Égert and Mihaljek (2007) point out that this equilibrium is not necessarily stable. In other literature, Dröes and van de Minne (2017) use a short-term error correction model to show that the time needed to absorb shocks out of equilibrium varied from 0 to 20 years, with average at 3 years. This suggest that the real estate market is more efficient in some periods than in others and that commonly used variables are not enough to capture the entire effect.

4 Literature review

In order to be able to detect and overcome turbulent periods, the main focus of researchers and central banks, especially after the last real estate bubble burst, is to study aggregate data for countries as a unit. Thus, the quantity of regional analysis is rather small in economic literature. Moreover, because of existing historic influence of planned economy and transition processes, not a lot of analysis of property price determinants has been done in CEE countries until recent years. Such studies also use data with shorter time periods than those of developed countries. For all these reasons, the results may differ substantially based on the dataset used in each study. This chapter presents papers which examine the determinants of real estate prices on an international level and within the Czech Republic. Studies which used data similar to those which are available for this paper will especially be considered. Furthermore, selected papers aiming to detect bubbles on the real estate market will be mentioned, as well as papers that inspired the analysis in this paper in terms of the econometric models they used.

The first paper that is used as a source of great inspiration is, to the authors' knowledge, the first comprehensive study of regional determinants of real estate prices in the Czech Republic so far, conducted by **Hlaváček and Komárek (2009b)**. Their study consists of two parts. The first objective of this study was to determine periods when real estate prices were overvalued and thus identify a bubble in the market. To do that, two different approaches were employed. The first approach uses price-to-rent and price-to-income ratios as standard indicators used in research of market bubbles. The second approach uses the Hodrick-Prescott filter with quarterly data of real estate prices. They reached a conclusion that the Czech real estate market experienced two bubbles in the examined time period, one in 2002/2003 and the second in 2007/2008. Although the increases in prices were similar in both periods, the price increase in 2007/2008 can be largely explained by fundamentals and therefore the level of overvaluation was lower in the 2007/2008 bubble.

In the second part, a panel regression analysis of apartment prices across the Czech regions is performed, using annual data for the Czech Republic, Prague and the Czech Republic without Prague in the period 1998-2008. Two alternative methods were used to analyse the determinants of property price - OLS with first differences of the apartment prices as the dependent variable and panel regression with fixed effects on the level of property prices. Authors also decided to estimate three different models, one that includes the full set of explanatory variables and one in which some of the variables (*building plot prices* and *monthly rent*) are excluded due to potential endogeneity problems. This second model was also estimated with Prague excluded as an outlier. All specifications reported similar results, which suggests that the endogeneity has not affected the estimation much. Demographic factors that proved to be significant are *divorce rate*, *natural population growth* and *net migration*. Out of labour market variables, only *unemployment rate* turned out to be important, which might relate to lower disposable income of households in regions with higher unemployment rates. As for other demand factors, the coefficients of *growth in market rent* and *growth of average monthly wage* were both significant and with the expected sign as well. What one might not expect was that the role of *housing loans* as a major demand determinant of property price was not confirmed. Authors explain this by exponential growth of housing loans in 2002-2008 irrespective of developments on the real estate market, but they expect the standard relationship to restore. Given that this is an unexpected result, this relationship will be tested in presented paper in a time period after the paper by Hlaváček and Komárek was published. Above that, the coefficient of *interest rates* was also insignificant and recorded with the opposite sign than was expected, which can be explained by the fact that interbank rates were used rather than rates for housing loans. The effects of supply factors were mixed as *building plot prices* were significant in both time series analysis and panel data regression. In some specifications, the *number of apartments per 1,000* *inhabitants*, was also significant. Other supply factors proved to have little or no affect on apartment prices.

Results of regional differences were as expected. Apartments in regions with lower prices were undervalued and apartments in regions with higher prices were overvalued. However, because of the specific nature of the capital city, apartment prices in Prague are *ceteris paribus* higher than in other legions. Since in the paper authors use annual data, the time series in their panel regression are quite short. By using quarterly data, the analysis conducted in this paper should be able to provide a more detailed view and also capture the seasonal effect.

One of the first detailed studies of real estate prices in the CEE region was carried out by **Égert and Mihaljek (2007)**. They used panel dynamic OLS with error correction model to study determinants of real estate prices in 19 developed OECD countries and 8 transition economies of Central and Eastern Europe (CEE). Their research question was whether the traditional fundamental determinants drive real estate prices in CEE similarly to OECD countries. They consider a set of standard demand and supply factors used in the empirical literature and some transition-specific factors, such as *institutional development* and *improvements in housing quality*. The analysis confirmed that these transition-specific variables have a fairly strong impact on property prices in CEE countries. Changes in *real interest rates* have a significantly higher impact on prices in the group of CEE countries. On the other hand, *credit growth* affects prices in OECD countries roughly two times more than in CEE countries. Fundamental factors, such as *GDP per capita*, *real interest rates*, *housing credit* and *demographic factors*, are highly significant in both CEE and OECD countries. Furthermore, *price elasticities* were observed generally higher in transition economies than in developed countries, which suggests that adjustment of property prices to the equilibrium is faster in CEE.

Čadil (2009) aimed to identify a bubble on the czech real estate market using the price-to-income ratio. Then he also used VAR analysis to find factors which determine the real growth of real estate prices from 1998 to 2006 using prices of both apartments and family houses. Using the Granger causality test, he found a possible dependence of the amount of mortgage loans on the apartment prices growth and reverse dependence of mortgages on the growth of family house price. In the VAR model, *average monthly wage* and *the share of population aged between 20-39 years* proved to be significant factors for the prices of apartments. In the case of family houses also the *real interest rate (1-year PRIBOR)*. A non-fundamental factor that proved to determine the price dynamic was the lagged value of the property price, or *speculative demand*. Its significance in this model may indicate the presence of a bubble in the czech housing market.

Zemčík (2009) used annual panel data for major cities of the Czech Republic from 2001 to 2008 and monthly data for major districts of Prague from july 2007 to february 2009 to analyze the relationship between *real estate prices* and *rents*. In order to identify areas with overvalued apartments, he employed the present value model with Granger causality techniques and price-to-rent ratio with stationarity tests, and concludes that bubbles are present predominantly in big cities, such as Prague, Olomouc and Hradec Králové.

Another recent study of real estate prices determinants was conducted by **Belke and Keil (2017)**. Authors of this study use annual data for the german regions (127 cities) in 1995-2009, with two dependent variables - *house prices* and *prices of newly built apartments*. Using the fixed effects panel regression, the authors were able to determine variables which proved to be robust determinants of real estate prices with their effects being in line with theoretical predictions. Supply factors, *the number of newly constructed apartments*, *the number of real estate transactions* and *the number of existing apartments*. On the demand side, *the number of households*, *quality of regional infrastructure* (measured by *the number of hospitals per 1,000 inhabitants*) and *number of people aged 15-65*, *number of households in each city*.

One of the recent studies of regional determinants in a transition country is **Cohen and Karpavičiūtė (2017)**. This paper investigated the impact of fundamentals on housing prices in Lithuania using quarterly data from 2001Q1 to 2014Q2. The Granger causality test showed that *GDP* and *unemployment* are causal determinants of *housing prices*, but there is no causal relation of *housing prices* with *interest rate* and *emigration*. However, a reverse relationship was found with *inflation*. This means that *housing prices* granger cause *inflation*. In this case, including *inflation* into a regression analysis as an explanatory variable could lead to incorrect results. For this reason, authors recommend to test causality of variables before running the regression.

5 Data and Methodology

This chapter provides a description of the dataset used in the empirical part of this paper as well as a discussion of issues related to data gathering.

5.1 Real Estate Prices in the Czech Republic

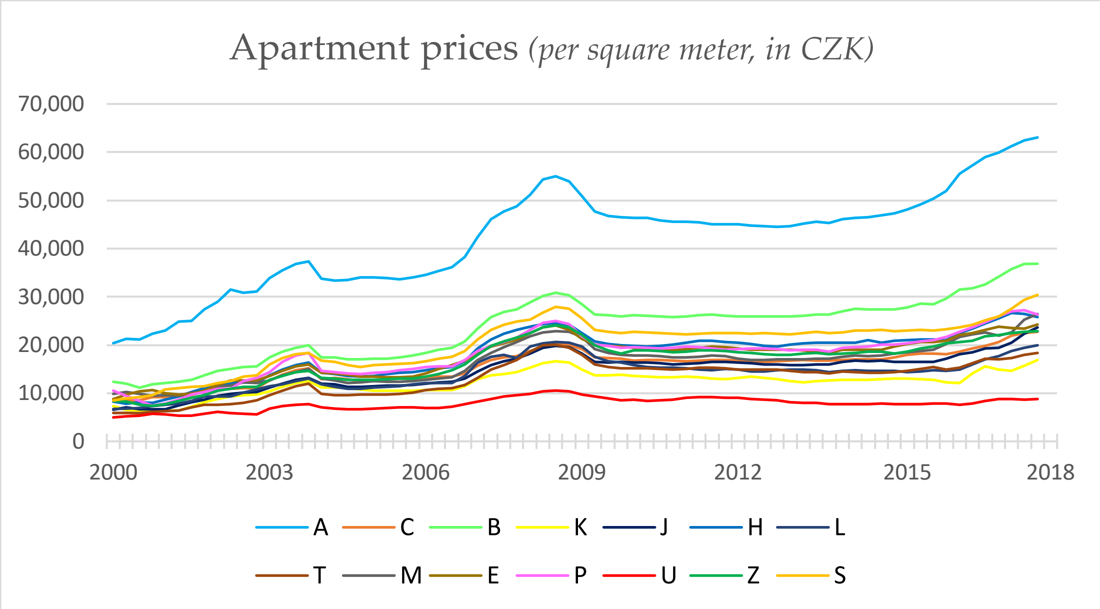
The core variable of interest in this analysis will be apartment price. In the Czech Republic, as a former transition economy, the real estate market was fully liberalized in the 1990's and the earliest publicly available data come from 1998. There are two types of data to choose from, based on the methodology of their gathering. Property transfer prices collect data from information on transactions and are therefore closest to the actual realized prices. Supply prices, on the other hand. are collected from bids of real estate agencies and their changes are likely to be distorted by different margins of individual agencies. (Hlaváček and Komárek 2009b)

The only data on transfer prices in the Czech Republic are collected by CZSO in "Monitored Type of Real Estate Prices". This publication collects data from statements for stamp duty land tax (SDLT). The advantage of this source is that it is based on real, actually paid prices. Since almost all transfers of second-hand apartments are subject to tax, it provides almost complete information on their prices. However, information about prices of new apartments are not included, as they are not subject to property transfer tax. Above that, these data cover all types of real estate and provide a classification based on the size of municipality and the degree of wear of a given type of real estate. All prices are calculated per square meter. The main disadvantage can be seen in the delay of approximately one year with which these data are published.

Supply prices are monitored by the Institute for Regional Information (IRI). Although the data series begins in 2000, it covers only yearly data until 2007. IRI also does not provide prices per square meter, but uses a "standard apartment" of 68 square meters and of a certain age. As a result, these data are limited to a small segment of the market and lack the completeness of the transfer prices. Although the data are collected for region and district municipalities, the IRI database is not public. Another source of supply prices is CZSO. However, it only monitors prices for Prague and the Czech Republic without Prague. Due to their clear methodology and complete regional and quarterly coverage, we decided to use *transfer prices* in this paper.

Although there are inherent differences between prices of individual apartments due to market features discussed in Chapter Two, apartments are considered more homogeneous than family houses or apartment buildings. Hlaváček and Komárek (2009b) also emphasize the advantage of apartment price dynamics, which are more sensitive than prices of other types of properties and more distinctly reflect the changes in underlying factors (see Figure A.1). Between 2015 and 2017, transfer of apartments accounted for 79% of all transfers of above mentioned types of real estate. Therefore, the number of transactions is another advantage of apartments. For these reasons, we decided to further work with the segment of *apartments* in the empirical part of this paper.

Figure 5.1: Apartment prices in Czech regions



Source: CZSO

The plot of apartment prices in Czech regions is available in Figure 5.1. One can notice a sharp increase in prices during the global financial crisis of 2007/2008 and subsequent fall in 2009. Although this pattern is followed by all Czech regions, it is most noticeable in Prague. On the other hand, apartment prices in Ústecký region (U) have been almost steady throughout the observed time period and during the crisis of 2007/2008 experienced little volatility. A similar peak to the one in 2007/2008 is observed in the year 2003. These two anomalies were identified as price bubbles by Hlaváček and Komárek (2009b) and Čadil (2009). Since 2015, prices have been increasing rapidly after almost 6 years of stagnation. In the fourth quarter of 2017. the average transfer price of apartments in the Czech Republic reached 27,122 CZK per square meter. In comparison, two years earlier in 2015Q4 the average price in CZ was 21,414 CZK per square meter. That means a 26.7% increase over the two-year period. Above that, the pattern of price movement is rather similar to the one in period 2007/2008. Figure A.2 depicts the estimated overvaluation of apartment prices in the period 2008-2017. An important piece of information for the purposes of this paper is that both the Prague (A) and Ústecký (U) regions react quite differently to shocks than other regions. Together with the fact that Prague has had the highest prices and overall economic performance in the observed time period, we can consider Prague as an outlier and keep this in mind in the data analysis.

5.2 Data Description

Similarly to transfer prices of apartments ("*apartment prices*"), most of the data on independent variables are available from publications of the Czech Statistical Office (CZSO) or its public database. Building plot prices (per square meter) can be found in "Monitored Type of Real Estate Prices". Population and demographic factors are published quarterly in "The Population of the Czech Republic". Data on the number of apartments, construction costs (per square meter of newly constructed apartments), average monthly wages and unemployment rate are available in CZSO's public database. Information on the volume of mortgage loans is provided by Ministry of Regional Development (MMR). Finally, the number of hospital beds, which will be used as a proxy for quality of public infrastructure, is supplied by Institute of Health Information and Statistics of the Czech Republic (UZIS). All above mentioned variables cover a time period from 2000Q1 to 2017Q4 for all Czech regions. Real effective exchange rate (*REER*), interest rates and rents (per square meter) are monitored by the Czech National Bank. REER is defined as the weighted average of Czech koruna relative to a basket of 13 strongest trade currencies deflated by the GDP deflator. While REER and interest rate are the same for all regions, rents are available for county seats only. As a result, there are no data on rents for the Středočeský region (S), because it has its county seat in Prague, which is a separate region. Above that, the usability of rents in regional analysis may be considered doubtful, as the data only include information for one municipality in each region.

Another variable which might be considered problematic is urbanization. Specifically, it measures the percentage of the population which lives in municipalities with over 20,000 inhabitants. Similarly to the number of hospital beds, urbanization represents a proxy for qualitative characteristics of a properties' surroundings. Apartments located in urban areas are exposed to externalities, both positive, such as easier commute, and negative, like pollution. Urban areas also experience lower unemployment rates and higher wages. For these reasons, the effect of urbanization on apartment prices is expected to be positive in all regions. However, because the Prague (A) region consists of only one city, the percentage of population living in urban areas is 100% in all time periods. Therefore, there is no reason to examine this variable for this region (A).

The complete list of variables used in this study and their summary statistics can be found in Table 5.1. The dataset covers a time period of 18 years (72 quarters) from the first quarter of 2000 to the fourth quarter of 2017. Combined with the panel size of 14 regions it makes a total of 1,008 observations. With the exception of rents as discussed above.

Table 5.1: Summary statistics

Variable	*Source*	*Obs.*	*Mean*	*St.Dev.*
Apartment price	CZSO	1008	17,801.02	9,113.66
Building plot price	CZSO	1008	2,114.81	1,379.32

New construction *(per 1,000 inhabitants)*	CZSO	1008	0.67	0.37
Existing apartments *(per 1,000 inhabitants)*	CZSO	1008	385.18	19.09
Construction cost	CZSO, CNB	1008	29,825.51	4,535.49
Wages	CZSO	1008	19,797.89	4,910.05
Unemployment	CZSO	1008	6.57	2.88
Marriages *(per 1,000 inhabitants)*	CZSO	1008	1.20	0.67
Divorces *(per 1,000 inhabitants)*	CZSO	1008	0.71	0.13
Natural growth of population *(per 1,000 inhabitants)*	CZSO	1008	-0.05	0.39
Net migration *(per 1,000 inhabitants)*	CZSO	1008	0.42	1.16
Share of population aged 15-64	CZSO	1008	69.34	1.99
REER	CNB	1008	101.03	10.05
Interest rate *(3M PRIBOR)*	CNB	1008	2.07	1.59
Mortgage loans	CNB	1008	1,833,337.00	2,570,152.00
Rent	CNB, IRI	936	95.19	24.18
Hospital beds	UZIS	1008	4,327.64	2,174.03
Urbanization	CZSO	1008	39.49	19.94

5.3 Preliminary Analysis

In order to be able to choose an appropriate model, tests for unit root and cointegration are applied to perform diagnostics for the data used in this analysis.

5.3.1 Stationarity

To begin with, the full range of available unit root tests have been applied for checking stability of the data along with the Hadri Lagrange multiplier test for checking stationarity. Using formal tests preceded a visual inspection of plots of individual variables. It is expected that price variables along with wages, rents and construction cost are non-stationary, as they have an upward trend (see Figures A.3, A.4 and A.5). However, if we remove the trend or include it in the regression model, then variables with a unit root can become trend stationary. In order to remove the stochastic trend altogether, one has to take a first difference of the data. Another possibility is to use a logarithmic transformation.

The first test considered in this section is the **Levin–Lin–Chu** (2002, LLC). The authors suggest that the statistic performs well when the number of cross sections (N) lies between 10 and 250 and there are 25 to 250 observations (T) per panel, combined with a sufficient condition of $N/T \rightarrow 0$. These assumptions fit the dataset quite well. Disadvantages of this test are that it assumes a common unit root process and it relies heavily on the assumption of cross-sectional independence. However, from the nature of the data, we can assume cross-sectional dependency, as all regions share a common government policy and are located in a relatively small area. Moreover, the LLC test is specified with the null hypothesis of unit root in all cross-sections against the alternative of stationarity in at least one cross-section, which is very restrictive and does not allow for some panels to follow a unit root process and some not. Results of the LLC test are available in Table B.1.

Similarly, the **Breitung** (2000) unit root test also works with common autoregression between regions. Although it assumes that $N \to \infty$ and $T \to \infty$, the author claims that it has a good test power in small samples as well. Table B.2 shows that the Breitung test found a unit root in all price variables, wages and rent, which is more consistent with our expectations than the results of LLC test.

Another widely used unit root test is the **Im–Pesaran–Shin** (2003, IPS), which allows for individual autoregressive process in each panel. This assumption seems to fit our regional dataset properly. Furthermore, this test also allows for serial correlation and heterogeneity of residuals. The null hypothesis is that all regions follow a unit root process against the alternative, which still allows a unit root for some, but not all, regions. Although, the test assumes standard normal distribution of the standartized t-bar statistic as $N \to \infty$, according to Monte Carlo simulations performed by Im et al (2003), the IPS test performs better than Levin–Lin–Chu in small samples. Table B.3 reveals that the IPS test supports our theory and shows that all prices, wages and rents follow a unit root process, along with the number of existing apartments, unemployment, age structure, mortgages and infrastructure.

The **Augmented Dickey-Fuller** (1979, ADF) test and **Phillips-Perron** (1988, PP) test arrive at similar results (see Tables B.4 and B.5). In the software used for this analysis (*Stata 12* and *EViews 9*), both are available with two specifications. First is the standard Fisher-type test, which follows asymptotic Chi-square distribution. Second specification is the **Choi** (2001) Z-statistic, which assumes asymptotic normality for $N \to \infty$. Barbieri (2006) performed Monte Carlo simulations which suggest that the Choi specification of ADF has better performance than other Fisher-type tests. Both versions are specified with null hypothesis of individual unit root process. The lag lenghts for all tests were chosen by *Hannah-Quinn information criterion* (HQC). The reason why we chose HQC will be covered in the next section of this paper.

Last but not least, the **Hadri** (2000) Lagrange multiplier test builds on the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test from time series testing. The null hypothesis is that all panels are stationary against the alternative of at least one panel following a unit root process. Although this straightforward hypothesis is very appealing, as the test can confirm the presence of a unit root, the performance of the Hadri LM test is quite poor. According to Hloušková and Wagner (2005), a strong serial correlation of the autoregressive or the moving average type can cause that the null hypothesis is incorrectly rejected. From the results of the Hadri LM test (see Table B.6), we can see that the null hypothesis of stationarity was rejected for all variables with the exception of marriages.

Based on the results of performed unit root tests, we can conclude that wages, unemployment rate, age structure, mortgages and rents are non-stationary, as this was reported by all of the tests listed above. Apart from the LLC test, all other tests established that apartment prices, building plot prices, existing apartments and infrastructure follow a unit root process as well. These results are consistent with our expectations and the economic theory. After performing the test on first differences of respective variables, we found out that their first differences are stationary. Therefore, we can conclude, that the level variables are integrated of order one, i.e. I(1). The results for construction costs are fairly mixed. However, it seems more likely that this variable is non-stationary as well.

Since there is quite a high variation among different data, we will also consider the natural logarithm of price variables. This transformation is commonly used in relevant literature. Specifically, we will take log of apartment prices, building plot prices, wages and rent. This will allow us to interpret their coefficients as a percentage change. All log variables were, therefore, tested for the presence of unit root by all tests listed above. All tests found a unit root process in log(rent) and all but LLC in log(wages), while their first-differenced versions are stationary. We can therefore safely assume that these variables are also I(1). The results of unit root tests for log(apartment prices) (or "*lapartprice*") and log(building plot prices) (or "*lplotprice*") are rather mixed. While both fisher-type tests rejected the null hypothesis of unit root for *lplotprice*, they failed to reject it for *lapartprice*. Choi-type tests present similar discrepancy. LLC and IPS tests rejected the null hypothesis of unit root for both variables. However, Breitung and Hadri tests arrived at the opposite result. That being said, IPS rejected the null hypothesis of unit root at only 10% significance level for *lapartprice*. Due to these inconclusive test results we are unable to confirm or reject the possibility of both variables being non-stationary. However, the tests suggest that the presence of a unit root process is more likely for *lapartprice*. Following the intuition behind used tests and the logical foundation for logarithmic transformation, one can assume that log did not remove the unit root process entirely and thus we will work with these variables as if they were non-stationary. As for other variables, such as exchange rate, interest rate and demographics, the unit root testing did not convincingly reject their stationarity. As there is no reason to believe otherwise, one can make the inference that they are indeed stationary.

5.3.2 Granger Causality

Before we proceed to cointegration testing, let us consider the causal relationships in our model. As real estate prices affect economic variables, we will apply Granger causality test to reduce the number of causal determinants of apartment prices. This method is based on analysis conducted by Cohen and Karpavičiūtė (2017) discussed in Chapter Four. The **Granger** (1969) causality test is defined as

$$(4.1) X_t = \sum_{j=1}^{m} \alpha_j X_{t-j} + \sum_{j=1}^{m} \beta_j Y_{t-j} + \varepsilon_t$$

$$Y_t = \sum_{j=1}^{m} \gamma_j X_{t-j} + \sum_{j=1}^{m} \delta_j Y_{t-j} + \eta_t$$

The null hypothesis for the test is that lagged values of X do not have explanatory power on variation in Y and *vice versa*. As the test assumes that all data are stationary, we will use the first difference of apartprice, plotprice, rent and mortgage, because they are I(1). On the other hand, new construction is stationary, so it will be used in level. These variables were selected based on the logic foundation provided in Chapter Three. In order to be able to choose the correct number of lags, a vector autoregression (VAR) model had to be estimated. Subsequently, both Schwarz and Hannah-Quinn information criteria determined that the optimal lag lenght is 4, which is the expected standard in quarterly data.

Table 5.2: Granger causality test

Null Hypothesis:	*Obs.*	*F-Stat.*	*Prob.*
building plot price does not Granger cause *apartment price*	938	5.0211	0.0005***
apartment price does not Granger cause *building plot price*		6.6150	0.0000***
rent does not Granger cause *apartment price*	871	1.1649	0.3249
apartment price does not Granger cause *rent*		15.6118	0.0000***
mortgage does not Granger cause *apartment price*	938	12.9487	0.0000***
apartment price does not Granger cause *mortgage*		14.4015	0.0000***
new construction does not Granger cause *apartment price*	938	1.2870	0.2733
apartment price does not Granger cause *new construction*		9.2614	0.0000***

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***

The test revealed that there are mutual causal relationships between the following pairs of variables: apartment prices - building plot prices and apartment prices - mortgages. The causality of building plot prices towards apartment prices is expected, as building plots are input for the construction of new apartments. The reverse effect can be explained by the fact that changes in apartment prices create changes in demand for new construction and thereby create pressure on prices of building plots. Similarly, the mutual interaction of apartment prices and the volume of mortgages was expected as apartment prices work as a collateral for mortgage loans. An interesting result is that apartment price Granger causes new construction and not *vice versa*. A possible explanation could be that the change of apartment prices lead to a change in demand which creates a pressure on construction companies, while the price of apartments is rather rigid and does not react to the number of newly built apartments. Furthermore, the Granger causality test showed that apartment prices cause rent. This is well consistent with our expectation that owners set the rental price to a certain level so that it yields a required percentage of value of the property. On the other hand, one can assume that rent does not create enough pressure on apartment prices to affect them.

According to the Granger causality test, neither rent nor new construction are causal determinants of apartment prices. And because this inference also has a solid logical foundation, both variables will be excluded from further analysis.

5.3.3 Cointegration

In Sections 5.3.1, we established that some of our variables are I(1), i.e. non-stationary. The next step is to find out whether or not there are any cointegrating relationships among them. A common approach in empirical literature is to take the first difference of the data and estimate the model using one of standard panel data estimators. However, Baltagi (2011) emphasizes that, although differencing non-stationary variables does capture the short-run dynamics, it destroys potential information about the long-term relationship between them. Furthermore, including these variables in the regression model in levels leads to spurious regression and thus incorrect results are obtained.

We begin with the **Kao** (1999) cointegration test. Kao follows the Engle-Granger (1987) two-step cointegration test, which is based on the analysis of residuals. In the first step, a regression using I(1) variables is performed, and in the second step it tests stationarity of residuals obtained in step one. If the variables are cointegrated, then the residuals are stationary, and *vice versa*. The Kao test requires cross-section parameters in the regression run in step one. As the cross-sections used in this paper are likely to be homogeneous, we expect the test to perform rather well. The number of lags included in the second stage regression was specified using *Akaike (1973) information criterion* (AIC), *Schwarz (1978) information criterion* (SIC) and *Hannah-Quinn (1979) information criterion* (HQC). Liew (2004) established that with a relatively large sample (120 or more observations), HQC outperforms other criteria. In constrast, AIC is a better choice in a small sample (60 observations and less). As the dataset used in this paper includes 1,008 observations, it can be inferred that HQC is the most suitable.

Kao test was performed on two sets of I(1) variables. The full set includes *log(apartment price)*, *log(building plot price)*, *existing apartments*, *construction cost*, *log(wages)*, *unemployment*, *age structure*, *mortgages* and *infrastructure*. However, because one can make the argument that some of the variables could be stationary, we decided to create a restricted set of variables by excluding *log(building plot price)* and *construction cost*. For both sets of variables, Kao test rejected the null hypothesis of no cointegration at 1% confidence level.

Table 5.3: Kao cointegration test

ADF	*Full set*		*Restricted set*	
	t-statistic	*prob.*	*t-statistic*	*prob.*
Schwarz IC	-4.5973	0***	-5.0559	0***

Hannah-Quinn IC	-4.5973	0***	-5.0559	0***
Akaike IC	-4.5973	0***	-5.0559	0***

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***

In order to confirm the results of the Kao test, the **Pedroni** (1999) cointegration test was employed. Pedroni is also residual-based and allows for heterogeneity. That is why it provides results for panel statistics and group statistics. Panel statistic (within-dimension) assumes common AR process and has alternative hypothesis that all panels are cointegrated, whereas group statistic (between-dimension), which assumes individual AR, uses the heterogeneous alternative of some cointegrating panels. Group statistics are the group mean extention of the panel version. Gutierrez (2003) proved that the group-rho statistic has the best power of them all. Furthermore, he also found that Kao test is more suitable than Pedroni test for homogeneous panel with low T. For Pedroni test we used the restricted set of variables, as the test does not work with more than 7 variables. Table 5.4 revealed that Pedroni test rejects the null hypothesis in all available specifications, which confirms the results of Kao test. Thus, it can be inferred that there is a strong cointegration present among I(1) variables.

Table 5.4: Pedroni cointegration test

Alternative Hypothesis	*Common AR coefs.*		*Individual AR coefs.*	
	Statistic	*Prob.*	*Statistic*	*Prob.*
v-statistic	3.0146	0.0013***		
rho-statistic	-4.1150	0***	-3.3323	0.0004***
PP-statistic	-6.8685	0***	-6.9458	0***
ADF-statistic	-4.7113	0***	-4.8441	0***

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***

Finally, the **Johansen** (1995) cointegration test was employed to find the number of cointegrated variables. Again, the test was used on both full and restricted sets of variables. Results (see Table 5.5) suggest that there are at least three cointegrating relationships among variables in the full set, as the Johansen test rejected the null hypothesis of at most 2 cointegrating relationships at 1% significance level. In the restricted set, the hypothesis of at most 2 cointegrations was rejected at only 10% significance level. Since it is customary to use the significance level of 5%, we can conclude that in the restricted set there are at least 2 cointegrating relationships.

Table 5.5: Johansen panel cointegration test

Hypothesis of no. of cointeg.	*Full set*		*Restricted set*	
	F-statistic	*Prob.*	*F-statistic*	*Prob.*
0	277.20	0***	177.40	0***

At most 1	149.10	0***	91.71	0***
At most 2	53.88	0.0023***	39.70	0.0702*
At most 3	31.91	0.2783	34.47	0.1857
At most 4	26.06	0.5700	32.56	0.2524

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***

6 Empirical Framework

For describing the long-run relations between cointegrated non-stationary variables in a panel dataset, there are two estimation methods available. One is the **fully modified OLS** (FMOLS) proposed by Phillips and Hansen (1990), which modifies the OLS estimator to make corrections for endogeneity and serial correlation, which arise from the presence of cointegrating relationship. The other estimator is the **dynamic OLS** (DOLS), which is attributed to Saikkonen (1991) and Stock and Watson (1993). DOLS eliminates the short-run correlations by adding lags and leads of first-differenced non-stationary explanatory variables in the OLS regression. While FMOLS is asymptotically unbiased, DOLS estimator is asymptotically efficient. Kao and Chiang (2000) performed Monte Carlo simulations, which arrived at the conclusion that DOLS is computationally simpler and performs better than FMOLS and standard OLS in estimating cointegrated panel regressions. Although both models are applicable, there is little empirical literature which uses FMOLS to estimate cointegrated regressions. Following the relevant literature, we will use the panel dynamic OLS (PDOLS) for the long-run relationship estimation. The model is described as follows:

$$Y_{i,t} = \alpha_i + \sum_{h=1}^{n} \beta_{i,h} X_{i,t} + \sum_{h=1}^{n} \sum_{j=-k_{i,1}}^{k_{i,2}} \gamma_{i,j} \Delta X_{i,t-j} + \varepsilon_{i,t} \quad (4.2)$$

where $Y_{i,t}$ is *log(apartment price)*, β's are estimated coefficient of explanatory variables X. The maximum lag and lead lenght will be determined by HQC as it is the most suitable for given dataset. Together with PDOLS, an error correction model (ECM) is employed to capture the short-run dynamics of apartment prices. ECM is specified as

$$\Delta Y_{i,t} = \alpha_i + \rho_i (Y_{i,t-1} - \sum_{h=1}^{n} \beta_{i,h} X_{i,t-1}) + \sum_{h=1}^{n} \gamma_{i,h} \Delta X_{i,h,t} + \varepsilon_{i,t} \quad (4.3)$$

where the coefficients of long run relationships in the second term are obtained by PDOLS. In this model, rho is the error correcting term (ECT). If this parameter has a negative sign and is statistically significant, it confirms that there exists a long run equilibrium. A positive ECT would imply that prices are not converging in the long run. The size of the parameter measures the speed of adjustment.

6.1 Model Specification

Due to a relatively large number of variables that enter our model, the estimation will be originally made on a restricted benchmark model including demand-side fundamentals - *wages*, *unemployment* and *age structure*. These variables were chosen because disposable income of households and size of labour force were established as important factors in determining prices of real estate in relevant empirical literature. Remaining explanatory variables will be added one by one to this baseline specification until an extended model can be identified. These variables will be selected based on their significance, correct sign and size of the coefficient. Furthermore, this has to be consistent with the theoretical foundation. This procedure is quite common in relevant literature and has been applied by Égert and Mihaljek (2007), Huynh-Olesen et al (2013), etc.

6.2 Regression Results

Table 6.1 presents the estimation results of panel regression models applied to the baseline specification, in which *log(apartment prices)* is regressed on *log(wages)*, *unemployment* and *age structure*.

Table 6.1: Panel regression results

log(apartment price)	*Pooled OLS*	*FE (1)*	*FE (2)*	*RE*	*ARDL*	*PDOLS (1)*
log(wages)	1.2101 (0.0387) ***	1.2374 (0.0210) ***	0.3201 (0.1291) **	1.2367 (0.0211) ***	0.8811 (0.0483) ***	0.8617 (0.0313) ***
unemployment	-0.0711 (0.0031) ***	-0.0214 (0.0022) ***	0.0071 (0.0016) ***	-0.0220 (0.0022) ***	-0.0695 (0.0070) ***	-0.0605 (0.0041) ***
age structure	0.0603 (0.0048) ***	0.0483 (0.0025) ***	0.0614 (0.0080) ***	0.0485 (0.0025) ***	0.0238 (0.0070) ***	0.0231 (0.0045) ***
constant	-5.9676 (0.6138) ***	-5.7356 (0.3362) ***	1.7176 (1.2772)	-5.7360 (0.3418) ***		
effects	-	fixed (one-way)	fixed (two-way)	random (one-way)	-	
adjusted R^2	0.71					0.77
overall R^2		0.65	0.38	0.65		
within R^2		0.85	0.96	0.85		
between R^2		0.62	0.01	0.62		
rho		0.85	0.98	0.73		

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***
Standard errors in parentheses

To start with, we compare the standard panel regression estimators with dynamic models, namely pooled OLS, fixed effects (FE), random effects (RE), autoregressive distributed lags (ARDL) and panel dynamic OLS (PDOLS). As for the

FE estimator, both one-way (1) and two-way (2) specification were used. While FE(1) controls for region-specific fixed effect, the FE(2) includes the time fixed effect as well. As we can see from the results, FE(2) yields less plausible coefficients and has worse fit than the one-way specification of FE estimator.

Next, the F-test was performed in order to compare one-way FE and pooled OLS. The null hypothesis of the test is that in the regression model

$$Y_{i,t} = \beta X_{i,t} + u_i + e_{i,t} \tag{4.4}$$

the fixed effects $u_i = 0$. As we can see in Table 6.2, the null hypothesis was rejected at 1% significance level. Hence, we can conclude, that non-zero region-specific fixed effects are present. FE is therefore superior to pooled OLS.

Table 6.2: F-test

F-test	*F(13, 991)*	*Prob>F*
H_0: all u_i=0	298.28	0***

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***

Similarly, Hausman test (see Table 6.3) confirmed that the individual-specific effects are fixed, not random. Another parameter to note is the rho-statistic, which shows the percentage of variation that is explained by individual-specific effects. A higher rho in FE also indicates the superiority of FE estimator. Therefore, we can conclude that, out of the standard panel regressions, the one-way fixed effects estimator was found out to perform the best with given dataset.

Table 6.3: Hausman specification test

Hausman test	*chi^2*	*Prob>chi^2*
H_0: difference in coefficients is not systematic	13.77	0.0032***

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***

In all baseline models, the coefficients of all three variables are significant at 5% level and have the expected sign. However, coefficient of log(wages) and age structure are much larger in all standard panel regressions, with the exception of two-way FE. This illustrates to what extent the results can be distorted by choosing the wrong model. As we are dealing with dynamic panel data, standard panel estimators are biased. The robustness of PDOLS is supported by the results of ARDL regression, which comes to very similar results. Although ARDL is more flexible with variables of different order of integration, dynamic OLS is used more often in empirical literature for estimating long run coefficients.

In the baseline model estimated by PDOLS, the coefficient of log(wages) indicates that a 1% increase in the average monthly wage would cause apartment prices to increase, on average, by 0.86%. Although the coefficient of unemployment seems to be much smaller, one has to keep in mind that this variable is used in level, not logarithm. Thus, it is interpreted as follows. If the unemployment rate decreased by one unit (one percentage point), the apartment prices rise by approximately 6%. Both these coefficient are consistent with the expectation, that the long-term relationship between disposable income of households and real estate prices is positive. Similarly, if the share of population aged 15-64 increases by 1 percentage point, the apartment prices increase by 2.3%, due to higher demand.

Table 6.4: Panel dynamic OLS regression results

log(apartment price)	*PDOLS (1)*	*PDOLS (2)*	*PDOLS (3)*	*PDOLS (4)*	*PDOLS (5)*	*PDOLS (6)*	*PDOLS (7)*
log(building plot price)		0.3796 (0.0342) ***					
existing apartments			0.0015 (0.0011)				
construction cost				0.0000 (0.0000) ***			
log(wages)	0.8617 (0.0313) ***	0.5898 (0.0382) ***	0.8163 (0.0560) ***	0.7243 (0.0274) ***	0.8832 (0.0278) ***	0.8624 (0.0337) ***	0.9071 (0.0308) ***
unemployment	-0.0605 (0.0041) ***	-0.0386 (0.0048) ***	-0.0602 (0.0047) ***	-0.0489 (0.0039) ***	-0.0622 (0.0041) ***	-0.0568 (0.0041) ***	-0.0524 (0.0043) ***
marriages					-0.0759 (0.0348) **		
divorces						-0.1333 (0.1210)	
natural growth of population							0.1710 (0.0343) ***
net migration							
age structure	0.0231 (0.0045) ***	0.0185 (0.0039) ***	0.0214 (0.0045) ***	0.0270 (0.0034) ***	0.0215 (0.0041) ***	0.0241 (0.0053) ***	0.0164 (0.0045) ***
REER							
interest rate							
mortgages							
infrastructure							
urbanization							
ECT	-0.04 (0.01) ***	-0.04 (0.01) ***	-0.06 (0.01) ***	-0.05 (0.01) ***	-0.04 (0.01) ***	-0.04 (0.01) ***	-0.03 (0.01) ***
adjusted R^2	0.77	0.85	0.84	0.88	0.80	0.79	0.78

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level*
Standard errors in parentheses

Table 6.4 (continued): Panel dynamic OLS regressions results

log(apartment price)	*PDOLS (8)*	*PDOLS (9)*	*PDOLS (10)*	*PDOLS (11)*	*PDOLS (12)*	*PDOLS (13)*	*PDOLS (14)*
log(building plot price)							0.3250

							(0.0354) ***
existing apartments							
construction cost							
log(wages)	0.9284 (0.0342) ***	0.8438 (0.0270) ***	0.8141 (0.0372) ***	0.7574 (0.0330) ***	0.8855 (0.0295) ***	0.8320 (0.0297) ***	0.6518 (0.0378) ***
unemployment	-0.0433 (0.0046) ***	-0.0517 (0.0038) ***	-0.0622 (0.0043) ***	-0.0550 (0.0036) ***	-0.0734 (0.0039) ***	-0.0496 (0.0046) ***	-0.0200 (0.0049) ***
marriages							0.0180 (0.0255)
divorces							
natural growth of population							-0.0488 (0.0376)
net migration	0.0488 (0.0100) ***						0.0430 (0.0076) ***
age structure	0.0116 (0.0050) **	0.0105 (0.0042) **	0.0306 (0.0056) ***	0.0361 (0.0047) ***	0.0170 (0.0042) ***	0.0279 (0.0044) ***	-0.0090 (0.0041)
REER		0.0102 (0.0012) ***					0.0099 (0.0014) ***
interest rate			-0.0170 (0.0093) *				
mortgages				0.0000 (0.0000) ***			
infrastructure					0.0001 (0.0000) ***		
urbanization						-0.0044 (0.0012) ***	
ECT	-0.04 (0.01) ***	-0.03 (0.01) ***	-0.04 (0.01) ***	-0.05 (0.01) ***	-0.04 (0.01) ***	-0.04 (0.01) ***	-0.03 (0.01) ***
adjusted R^2	0.82	0.80	0.82	0.87	0.85	0.79	0.85

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***
Standard errors in parentheses

In order to be able to choose the best specification of the panel regression model, each variable was added one by one to the baseline specification PDOLS(1) in Table 6.4. Results of these regressions are denoted PDOLS(2) - PDOLS(13).

Firstly, supply factors are added in models (2), (3) and (4). While the impact of building plot prices on apartment prices is significant at 1% significance level, the effects of construction cost and number of existing apartments is negligible. In case of construction cost, there are differences in methodology compared to apartment prices. In order to calculate construction costs, CZSO uses utility floor space. Apartment prices are calculated using living area floor space. Moreover, construction cost measures the cost of newly constructed apartments, which are inherently more expensive than second-hand apartments. Therefore, these two variables are not easily comparable. On the other hand, building plot price has the correct positive sign and is statistically significant. Thus, it will be included in the extended regression model. As Granger causality test in Section 5.3.2 revealed, there is a mutual causal relation between prices of apartments and building plots. This can be explained by the substitutive relationship between these two assets as investment opportunities. As discussed in Section 3.1, the growth of apartment prices generates pressure on building plot prices via higher demand. Nevertheless, a significant long-run coefficient of building plot price confirms the results of the Granger causality test and we can make the inference that building plot prices do affect apartment prices in the long run.

Models (5) through (8) examine the effects of demographic factors on apartment prices. An interesting outcome of model (5) is that the long-run coefficient of marriage rate is negative. It suggests that with an increase in the number of marriages per 1,000 inhabitants, the apartment price decreases by 7.6%. This can be explained by the fact that married couples live together in one household before marriage or they live in two separate households and after marriage they move together, by which the demand for housing decreases. On the other hand, divorce rate turned out to be insignificant. Granger causality test (see Table B.7) also proved that divorce rate does not Granger cause apartment price. One possible explanation is that after a divorce, although one household is separated in two, both individuals are unable to purchase a dwelling very shortly afterwards, but rather decide to rent one. Thus, the demand for housing is not affected. Lastly, the coefficients of natural population growth and net migration are both positive and significant at 1% level. As these results support our expectations, both variables will be included in the extended model.

Two region-invariant factors (REER and interest rate) are evaluated in models (9) and (10). Firstly, the results of panel regression (9) show that the coefficient of REER is positive and statistically significant at 1% level. The positive sign of the coefficient implies that each loss in trade competitiveness leads to an increase of apartment price. Although the effect is rather small, the sign of the coefficient goes against economic theory. Mahalik and Mallick (2016) argue that there may be a lot of uncertainties around the returns of investments in real estate caused by the volatility in exchange rate, which may negatively affect the interest of foreign investors. Model (10) shows that an increase of one percentage point in interest rate leads to a 1.7% decrease in the prices of apartments. The correct sign of the coefficient implies that increases in the interest rate make loan financing the purchase of a property less attractive and the demand declines. However, because it is only significant at 10% level, this variable will not be included in the extended model in order to make it as parsimonious as possible.

Lastly, let us examine the effects of the number of mortgage loans, quality of public infrastrucure and urbanization in models (11), (12) and (13). Although all these variables are statistically significant, their coefficients are too small to have any noticible effect on apartment prices. For instance, the coefficient of urbanization can be interpreted as follows. If urbanization increased by one percentage point, the apartment price would decrease by 0.44%. However, as urbanization, on average, experiences changes of less than 0.5 percentage point per year, its overall effect is of little importance. As for mortgages and infrastructure, their estimated effects are even smaller. As a result, neither of these factors will be included in the extended regression model.

In order to capture the short-term dynamics of apartment prices, an error correcting term (ECT) was included in Table 6.4. Each ECT was estimated using lagged residuals obtained from given PDOLS regression. As we can see in models (1) through (13), the ECT has the correct sign and is significant at 1% level. This implies that the process (*apartment price*) is converging to a long-term equilibrium after short-term shocks. The size of the coefficient has values between -0.03 and -0.06, which suggests that in every time period, i.e. every quarter, there is between 3% and 6% adjustment towards the equilibrium.

Column (14) of Table 6.4 shows the results of the extended model regression estimated by panel dynamic OLS. The following variables were included based on the model selection procedure described in Section 6.1 - building plot price, wages, unemployment, marriages, natural increase of population, net migration, age structure and REER. We can see that some of the included variables have lost their statistical significance. This implies that demographic factors (marriages, natural growth of population and age structure) are less important than others. All other variables are still significant at 1% level and have the same sign as in their respective reduced models. One can notice that the size of some coefficients has decreased. These changes are likely to happen when estimating a more complex regression model, as the uncertainty increases with the number of variables. Nevertheless, we can conclude that the demand-side factors (wages and unemployment) are important determinants of apartment prices. Out of supply side factors, only building plot prices proved to be important. Although the importance of demographic factors has been disputed by the extended model, we can infer that they play a role in determining apartment prices. Moreover, net migration proved to be significant even in the extended model. Finally, the positive effect of REER on apartment prices has been confirmed in the extended regression. Furthermore, this analysis has shown that other factors are of low importance in determining apartment prices.

6.3 Regional Analysis

In this part of the paper, the extended panel regression model PDOLS(14) from Section 6.2 is used to examine different effects of housing price determinants across regions. From the original dataset, which includes data for all 14 Czech regions, we will exclude two groups of regions and run the regression on each group separately. This will enable us to check robustness of the extended model.

Prague (A) has been identified as a "*high-income*" outlier for several reasons. Firstly, important indicators used in this analysis, such as apartment prices, building plot prices, rents, etc., are far higher for Prague than for other regions. Secondly, as the seat of government and the Czech hub of international businesses, Prague's labour market is quite specific as well. High wages and low unemployment rate are accompanied by high GDP and, unlike most of Czech regions, a labour inflow. **Středočeský region (S)**, which surrounds Prague, is likely to share some of Prague's specific qualities. This applies particularly to districts Prague-East and Prague-West, which are directly adjacent to the Prague region. For that reason, this analysis would be best done on a district level, rather than region level. Unfortunately, for most variables, the data on district levels are not available. Therefore, we will estimate two models - one in which Prague is excluded and another, where both Prague and Středočeský region are excluded.

The second group of regions which differ from the rest are **Moravskoslezský (T)** and **Ústecký (U) regions**. In the 20th century, both were highly industrialized regions. Local coal mines supplied thousands of jobs and large apartment building were provided for miners. However, in 1990's the mining industry declined dramatically as coal mining was no longer profitable after the transformation of Czech economy. As a result, thousands of workers were laid off and the unemployment rate increased rapidly. Furthermore, many apartments were abandoned and their prices declined. In the observed time period of 2000Q1-2017Q4, both regions experienced high unemployment rates, low apartment prices and persistent population outflow, i.e. negative net migration. Due to the specific nature of demographic and labour factors in these regions, both will be excluded from our dataset.

The results of regional analysis are revealed in Table 6.5, which is divided into three parts. The first three columns show results for models, where Prague (A) is analyzed separately in column (1) by dynamic OLS, the model for Moravskoslezský (T) and Ústecký (U) regions is estimated by panel dynamic OLS in column (2) and the rest of Czech Republic (11 regions) is also analyzed by panel dynamic OLS in column (3). Columns (4) through (6) show results for models where Středočeský region (S) is excluded from CZ and is estimated together with Prague (A) in column (4). In order to provide a direct comparison, the results of panel regression for all 14 regions conducted in Section 6.2 is provided in column (7).

Table 6.5: Panel dynamic OLS regression results by region

log(apartment price)	*A* *(1)*	*T+U* *(2)*	*rest of CZ* *(3)*	*A+S* *(4)*	*T+U* *(5)*	*rest of CZ* *(6)*	*CZ* *(7)*
log(building plot prices)	0.6755 (0.2931) **	-0.8987 (0.1944) ***	0.0286 (0.0408)	0.5155 (0.0560) ***	-0.8987 (0.1944) ***	0.0302 (0.0419)	0.3250 (0.0354) ***

log(wages)	0.4935 (0.2422) **	1.0969 (0.1918) ***	0.9160 (0.0409) ***	0.5303 (0.0590) ***	1.0969 (0.1918) ***	0.9228 (0.0441) ***	0.6518 (0.0378) ***
unemployment	-0.0190 (0.0271)	-0.0412 (0.0155) ***	-0.0101 (0.0052) *	-0.0397 (0.0163) **	-0.0412 (0.0155) ***	-0.0076 (0.0055)	-0.0200 (0.0049) ***
marriages	0.0745 (0.0330) **	-0.0099 (0.0745)	0.0082 (0.0270)	0.0543 (0.0391)	-0.0099 (0.0745)	0.0113 (0.0328)	0.0180 (0.0255)
natural growth of population	-0.2497 (0.0824) ***	-0.0795 (0.1373)	-0.0372 (0.0409)	-0.1634 (0.0642) **	-0.0795 (0.1373)	-0.0417 (0.0456)	-0.0488 (0.0376)
net migration	0.0127 (0.0097)	-0.0413 (0.0389)	0.0566 (0.0092) ***	0.0203 (0.0103) *	-0.0413 (0.0389)	0.0828 (0.0157) ***	0.0430 (0.0076) ***
age structure	-0.0129 (0.0080)	0.0564 (0.0173) ***	-0.0100 (0.0043) **	0.0031 (0.0083)	0.0564 (0.0173) ***	-0.0112 (0.0047) **	-0.0090 (0.0041)
REER	0.0059 (0.0029) **	0.0098 (0.0037) ***	0.0122 (0.0014) ***	0.0056 (0.0033) *	0.0098 (0.0037) ***	0.0120 (0.0015) ***	0.0099 (0.0014) ***
ECT	-0.17 (0.07) **	-0.09 (0.03) ***	-0.05 (0.01) ***	-0.14 (0.05) ***	-0.09 (0.03) ***	-0.05 (0.02) ***	-0.03 (0.01) ***
adjusted R^2	0.94	0.86	0.86	0.98	0.86	0.85	0.85

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***
Standard errors in parentheses

An interesting outcome of the first three models is that building plot price plays a very different role in each group of regions. While in Prague building plot price has the expected positive effect on apartment price, in the low-income regions (T+U), the relationship is negative. Furthermore, in the model for the rest of Czech regions, this variable was statistically insignificant. Due to the increasing shortage of buildable land in Prague, its availability together with building plot prices are expected to have a large impact on apartment prices. On the other hand, there have been thousands of disused apartments available in T+U regions in the early 2000's. Since the dependent variable is the price of used apartments, so it can be assumed that the price increases as empty apartments are being purchased, while the building plots are in substitutive market. Similar results are reported in models (4) through (6), which confirms our hypothesis.

The effects of unemployment rate are rather mixed and difficult to draw any conclusions from, as the test results may be influenced by the different number of cross-sections used for each model. For instance, the model for Prague does not consider the cross-section dimension at all. Therefore, it fails to reflect the cross-sectional dependency of the data. This inconsistency must be considered when interpreting results. The results suggest that wages have smaller impact on apartment prices in Prague (A) and Středočeský region (S) than other regions. This can be explained by higher demand in both regions coming from abroad and from investors, whose purchasing power is not purely determined by local wages. This may also be the reason for the fact that the age structure, which provides information about the labour force, is statistically insignificant in these regions.

Demographic factors seem to have rather uneven effects on apartment prices in each group of regions. Although the marriage rate does have a positive coefficient in Prague, it is statistically insignificant in all other models. This confirms our findings from Section 6.2 that it is not an important factor in determining apartment prices. A similar inference can be made about the natural growth of population, although it has an unexpected negative coefficient in models (1) and (4). However, it is interesting to note that the effect of net migration is positive and its coefficient is significant at 5% level only in models which exclude both high-income and low-income regions. The different results for demographic factors in excluded regions can be explained by their unique labour market characteristics mentioned above. On the other hand, all results for models (3) and (6) are consistent with the model for CZ, which demonstrates the robustness of the regression model. Last but not least, we can notice that results for REER vary only marginally, as REER is the same for all regions.

To the authors' knowledge, this is the first regional analysis of housing price determinants in the Czech Republic, which uses panel regression techniques on groups of regions. Cempírek (2014) used residual based approach to describe price misalignments on three groups of regions - Prague (A), Moravskoslezský and Ústecký region (T+U) and the rest of the country (11 regions). Although this analysis presents useful insights on trends and stability of housing prices, there is no information provided about the different impacts of housing price determinants across regions.

6.4 Impact of the Global Financial Crisis

The global financial crisis of 2007/2008 has shown that real estate prices have a large impact on financial institutions. In order to preserve the macroeconomic stability, the study of housing price determinants is essential for central banks. This section of the paper describes how the relationships between apartment price and its determinants was changed by the events of 2007/2008. This will enable us to observe their relations in different parts of the real estate cycle. As apartment prices peaked in the second half of 2008, our original dataset will be divided into two time periods. The first half covers a period from 2000Q1 to 2008Q4, while the second half starts at 2009Q1 and ends at 2017Q4. Finally, the extended regression model from Section 6.2 was applied to both datasets. Their results in Table B.8 are accompanied by the regression results from Section 6.2 for easy comparison.

Although analyzing determinants of housing prices on datasets divided by time period is not common in the empirical literature, there are several studies conducted outside of the Czech Republic. For example, Benamraoui (2010) performed a multiple regression analysis of housing prices in the United Kingdom, which compared periods of 01/2002-10/2007 and 11/2007-06/2009. Following this format, we decided to further divide the dataset into periods in which the apartment price follows a certain pattern. Firstly, there are two periods in which apartment prices are overvalued (bubble of 2003/2004 and financial crisis of 2007/2008). Both periods are followed by the stagnation of prices (2005Q1-2006Q4 and 2009Q1-2014Q4). Last but not least, two more periods of a price growth are included (2000Q1-2002Q4 and 2015Q1-2017Q4).

The results in Tables B.8 and B.9 show that the effect of building plot prices has been higher after the financial crisis of 2008 and the effect of wages has been smaller. In the case of building plot prices, they experienced a steep growth during the financial crisis of 2008 only in Prague region (A), while the dynamic in other regions was generally unaffected by the crisis. A more interesting finding is that the coefficient of unemployment is highest during the financial crisis, which is when the unemployment rate experienced a sudden rise in all regions (see Figure A.6). When looking at the coefficients of demographic factors, we can see that their signs and magnitudes are mixed and quite difficult to interpret. However, we can observe that the effect of net migration is the most stable, as it has the expected positive sign and is significant at 5% level in most periods. The coefficients of age structure have the same negative sign in both models. They are significant in only three time periods. Furthermore, we are unable to draw any conclusions from the results of REER as they vary quite a lot in all time periods.

As the estimation of long-term relationships between variables are less informative for data with short time periods, both models are merely an extension of the analysis in previous chapters and their results should be interpreted with caution. Furthermore, it must be stressed that this is the first analysis of housing price determinants in Czech Republic, which focuses on the change of their dynamics in different time periods, including the effect of the global financial crisis. This area of study may be an interesting subject of further research.

7 Conclusion

Due to the complexity of the real estate market, housing price developments are influenced by many variables. The goal of this paper was to identify the most important factors which determine the prices of residential real estate in Czech regions. Several methods were used in order to analyze the effects of explanatory variables on apartment prices. Traditional panel regression models, such as pooled OLS, fixed effects and random effects, were compared to dynamic panel regression models, namely autoregressive distributed lags and dynamic OLS. Based on the nature of the available data, an extended panel dynamic OLS model was identified in order to describe the long-term equilibrium of the relationship between apartment prices and their determinants. Furthermore, an error correction model was employed in order to capture the short-run dynamics of apartment prices.

The PDOLS regression revealed that apartment prices in the Czech Republic are driven mainly by demand-side fundamentals: wages and unemployment. These results are consistent with empirical literature related to this paper, namely Hlaváček and Komárek (2009b) and Čadil (2009). Of the supply factors, only building plot prices turned out to be significant. Furthermore, the estimated long-run effect of this variable was larger than in previous studies. This finding adds valuable information to the existing literature, as it was supported by several robustness checks. Although the effects of the demographic factors were mixed, a positive effect of net migration on apartment prices was established. The regression also showed that there is an unexpected positive relationship between REER and apartment prices. Looking at the results of the error correction model, one can see that the coefficients of ECT are negative, which implies that the apartment price converges to its long-term equilibrium after shocks. In addition to this, the Granger causality test showed that the apartment price Granger causes rents and new construction, not *vice versa*.

The robustness of the extended regression model was verified by excluding outlier regions and examining them separately. The results revealed that the effect of building plot price on apartment price is negative in Moravskoslezský (T) and Ústecký (U) regions. On the other hand, labour force factors (wages and unemployment rate) seem to be more important in these regions than they are in the high-income region of Prague. Furthermore, the regression established that the results for Prague are very similar to the results for both Prague (A) and Středočeský (S) regions when estimated together.

The last section of the paper examines the developments of relations between apartment price and its determinants by time period. We found that the impact of building plot price has become more important after the global financial crisis, while the effect of wages on apartment price has decreased. Furthermore, the effect of unemployment rate peaked in the period of 2007-2008, which supports the economic theory. The changes of other relations are rather mixed.

To the best of the authors' knowledge, this is the first application of panel data regression of real estate prices by region and by time period. The use of extensive robustness checks is therefore particularly innovative for the literature examining housing price determinants in the Czech Republic. The presented regressions could be further enhanced by analyzing each region individually. However, in order to provide significant results, it requires a longer time series of the data. Similarly, an extension to regions in multiple countries of Central Europe would put the results of this study in a broader perspective.

References

Al-Marwani, H.A. (2014): "An Approach to Modeling and Forecasting Real Estate Residential Property Market", Doctoral Thesis, Brunel University, United Kingdom

Apergis, N. & Rezitis, A. (2003): "Housing Prices and Macroeconomic Factors in Greece: Prospects within the EMU", *Applied Economics Letters*, vol. 10(9)

Baltagi, B.H. (2011): "Econometrics", 5th edition, *Springer*, Berlin, Germany

Barbieri, L. (2006): "Panel unit root tests: A review", Quaderni del Dipartimento di Scienze Economiche e Sociali, Serie Rossa, vol. 43, Università Cattolica del Sacro Cuore, Piacenza

Belke, A. & Keil, J. (2017): "Fundamental determinants of real estate prices: A panel study of German regions", *Ruhr Economic Papers*, vol. 731, Leibniz-Institut für Wirtschaftsforschung, Essen, Germany

Benamraoui, A. (2010): "The Real Effect of the Recent Financial Crisis on the UK Housing Price Indicators", *International Journal of Arts and Sciences*, vol. 3(13), pp. 255-266

Brereton, F. & Clinch, P.J. & Ferreira, S. (2008): "Happiness, geography and the environment", *Ecological Economics*, *Elsevier*, vol. 65(2), pp. 386-396

Capozza, D.R. & Hendershott, P.H. & Mack, C. & Mayer, C.J. (2002): "Determinants of Real House Price Dynamics", *NBER Working Papers*, no. 9262, October

Cohen, V. & Karpavičiūtė, L. (2017): "The Analysis of the Determinants of Housing Prices", *Independent Journal of Management & Production (IJM&P)*, vol. 8(1), pp. 49-63

Čadil, J. (2009): "Housing Price Bubble Analysis - Case of the Czech Republic", *Prague Economic Papers*, 2009, vol. 2009(1), pp. 38-47

CZSO (2014): "Domovní a bytový fond podle výsledků sčítání lidu (Housing and Dwelling Stock by Population and Housing Census)"

Do, A. & Grudnitski, G. (1997): "The impact on housing values of restrictions on rights of ownership: The case of an occupant's age", *Real Estate Economics*, vol. 25(4), pp. 683-693

Dröes, M.I. & van de Minne, A. (2017): "Do the Determinants of House Prices Change over Time? Evidence from 200 Years of Transactions Data", *European Real Estate Society (ERES)*, Presented at: ASSA-AREUEA 2017, Chicago, USA

Égert, B. & Mihaljek, D. (2007): "Determinants of House Prices in Central and Eastern Europe", *Comparative Economic Studies*, october 2007, vol. 49(3), pp. 367–388

Forgey, F.A. & Rutherford, R.C. & VanBuskirk, M.L. (1994): "Effect of Foreclosure Status on Residential Selling Price", *Journal of Real Estate Research*, *American Real Estate Society*, vol. 9(3), pp. 313-318

Fujiwara, D. & Campbell, R. (2011): "Valuation Techniques for Social Cost-Benefit Analysis: Stated Preference, Revealed Preference and Subjective Well-Being Approaches", HM Treasury, London, United Kingdom

Gutierrez, L. (2003): "On the Power of Panel Cointegration Test: A Monte Carlo Comparison", *Economics Letters*, vol. 80(1), pp. 105-111

Hilbers, P. & Hoffmaister, A.W. & Banerji, A. & Shi, H. (2008): "House Price Developments in Europe: A Comparison", *IMF Working Papers*, vol. 08/211

Hlaváček, M. & Komárek, L. (2009a): "Property Price Determinants in the Czech Regions", *CNB Financial Stability Report 2008/2009*, pp. 82-91

Hlaváček, M. & Komárek, L. (2009b): "Housing Price Bubbles and their Determinants in the Czech Republic and its Regions", *CNB Working Papers*, vol. 12/2009

Hloušková, J. & Wagner, M. (2005): "The Performance of Panel Unit Root and Stationarity Tests: Results from a Large Scale Simulation Study", *Discussion papers*, march 2005, Universität Bern, Switzerland

Huynh-Olesen, D.T. & Steiner, K. & Hildebrandt, A. & Wagner, K. (2013): "Residential Property Prices in Central, Eastern and Southeastern European Countries: The Role of Fundamentals and Transition-Specific Factors" *Focus on European Economic Integration*, Oesterreichische Nationalbank (Austrian Central Bank), Q2/2013, vol. 2, pp. 52-76

Im, K.S. & Pesaran, M.H. & Shin, Y. (2003): "Testing for Unit Roots in Heterogeneous Panels", *Journal of Econometrics*, july 2003, vol. 115(1), pp. 53-74

Ioannides, Y.M. & Rosenthal, S.S. (1994): "Estimating the Consumption and Investment Demands for Housing and Their Effect on Housing Tenure Status", *The Review of Economics and Statistics*, vol. 76(1), pp. 127-141

Kao, Ch. & Chiang, M.H. (2000): "On the estimation and inference of a cointegrated regression in panel data", *Nonstationary Panels, Panel Cointegration and Dynamic Panels*, vol. 15, pp. 179-222

Láznička, J. (2016): "Impact of metro station proximity on apartment value in Prague", Bachelor Thesis, Charles University, Czech Republic

Li, L. (2015): "Prices and Bubbles: Factors Affecting the Chinese Real Estate Market", Master's Thesis, Lund University, Sweden

Liew, V.K.S. (2004): "Which Lag Length Selection Criteria Should We Employ?" *Economics Bulletin*, vol. 3(33), pp. 1–9

Lux, M. (2009): "Housing policy and housing finance in the Czech Republic during transition: An example of the schism between the still-living past and the need of reform", *Sustainable Urban Areas*, Delft University of Technology

Mahalik, M.K. & Mallick, H. (2016): "Are house prices guided by fundamentals or speculative factors? An empirical inquiry for India", *International Journal of Economic Policy in Emerging Economies*, vol. 9(1), pp. 47-64

Maher, A. (2018): "Long-run drivers and short-term dynamics of Swedish real house prices", *International Journal of Housing Markets and Analysis*, vol. 11(1), pp. 45-72

OECD (2011): "Doing Better for Families", *OECD Publishing*, Paris, France, <https://doi.org/10.1787/9789264098732-en>

Reichel, V. & Zimčík, P. (2018): "Determinants of Real Estate Prices in the Statutory of Brno", *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, vol. 66(4), pp. 991-999

Sirmans, S.G. & MacDonald, L. & Macpherson, D.A. & Zietz, E.N. (2006): "The Value of Housing Characteristics: A Meta Analysis", *Journal of Real Estate Finance and Economics*, 11, vol. 33(3), pp. 215-240

Steel, M.F.J. (2017): "Model Averaging and its Use in Economics", *MPRA Paper 81568*, University Library of Munich, Germany

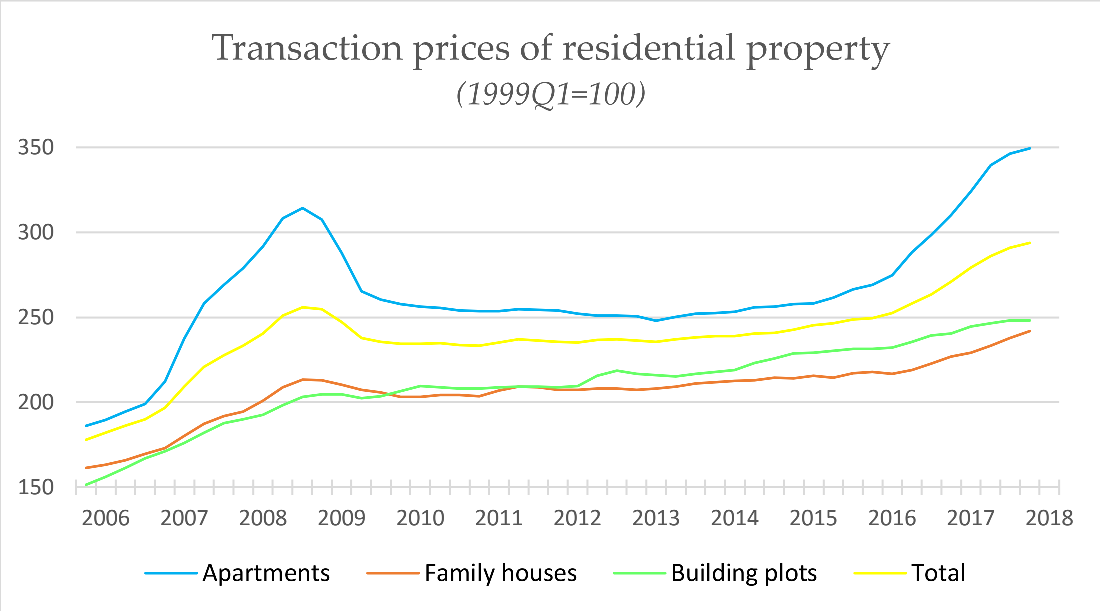
Tsatsaronis, K. & Zhu, H. (2004): "What drives housing price dynamics: cross-country evidence", *BIS Quarterly Review*, March

Zemčík, P. (2009): "Housing markets in Central and Eastern Europe: Is There a Bubble in the Czech Republic?", *CERGE-EI Working papers*, volume 390, September

Yusupova, A. (2016): "An Econometric Analysis of U.K. Regional Real Estate Markets", Doctoral Thesis, Lancaster University, United Kingdom

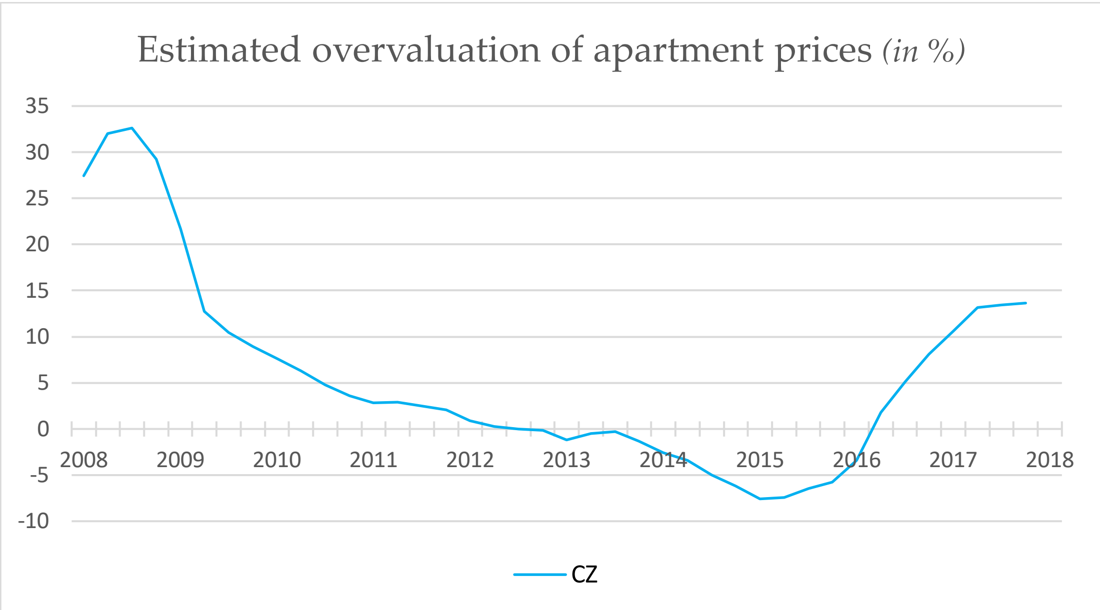
Appendix A: Figures

Figure A.1: Transaction prices of residential property in the Czech Republic



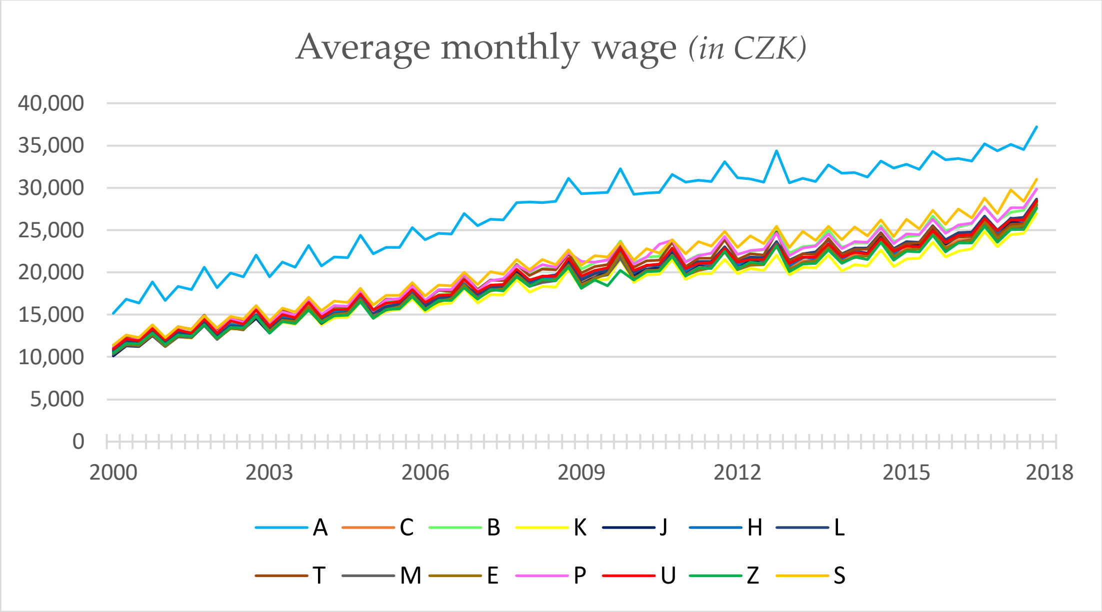
Source: CZSO

Figure A.2: Estimated overvaluation of apartment prices



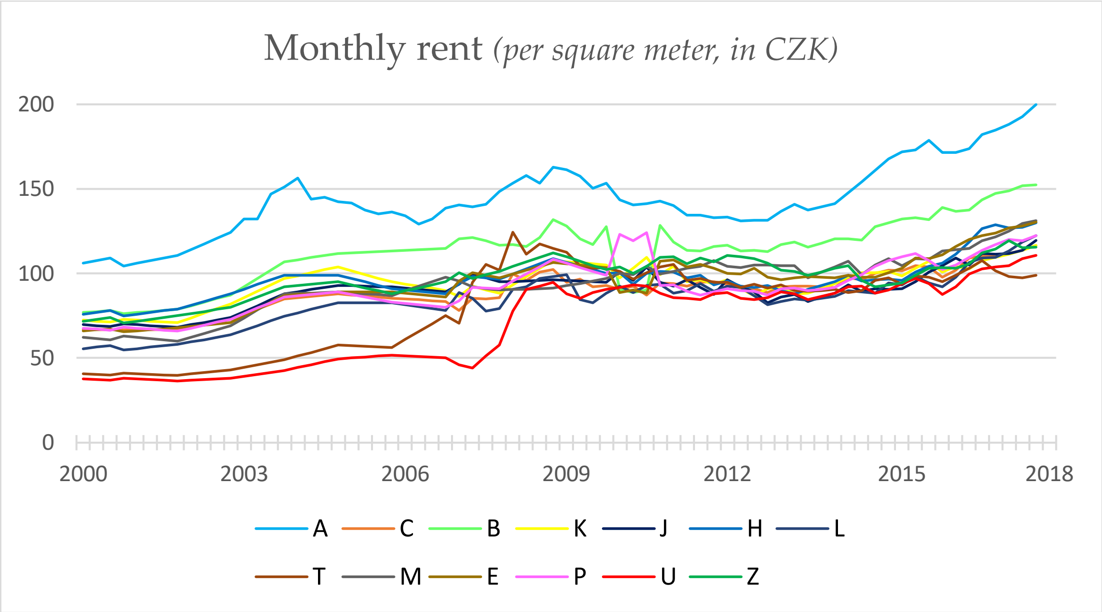
Source: CNB Financial Stability Report 2017/2018

Figure A.3: Average monthly wage in Czech regions



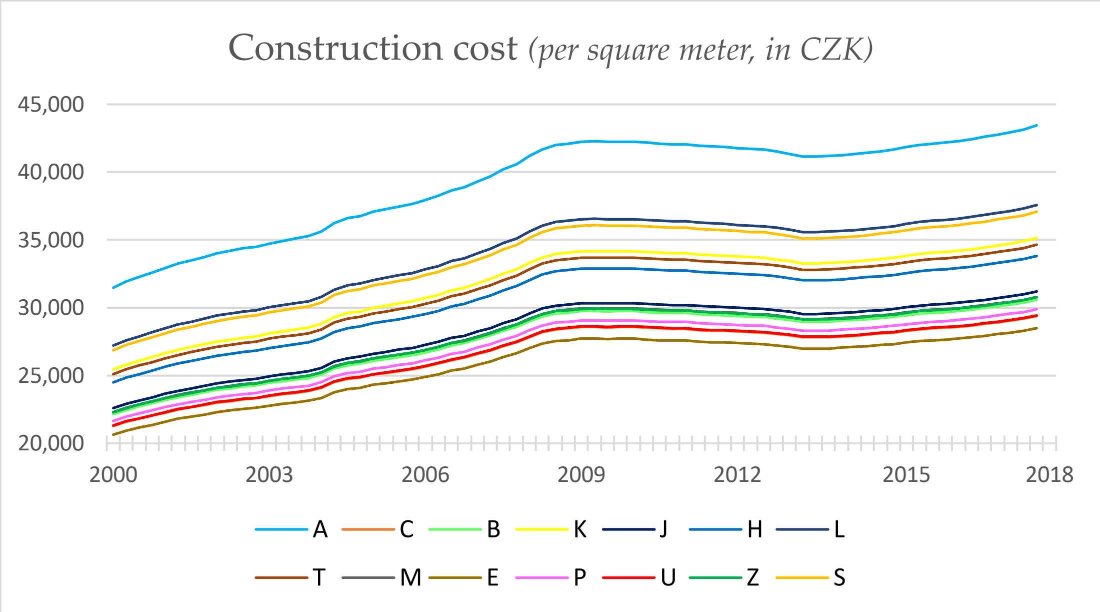
Source: CZSO

Figure A.4: Monthly rent in Czech regions



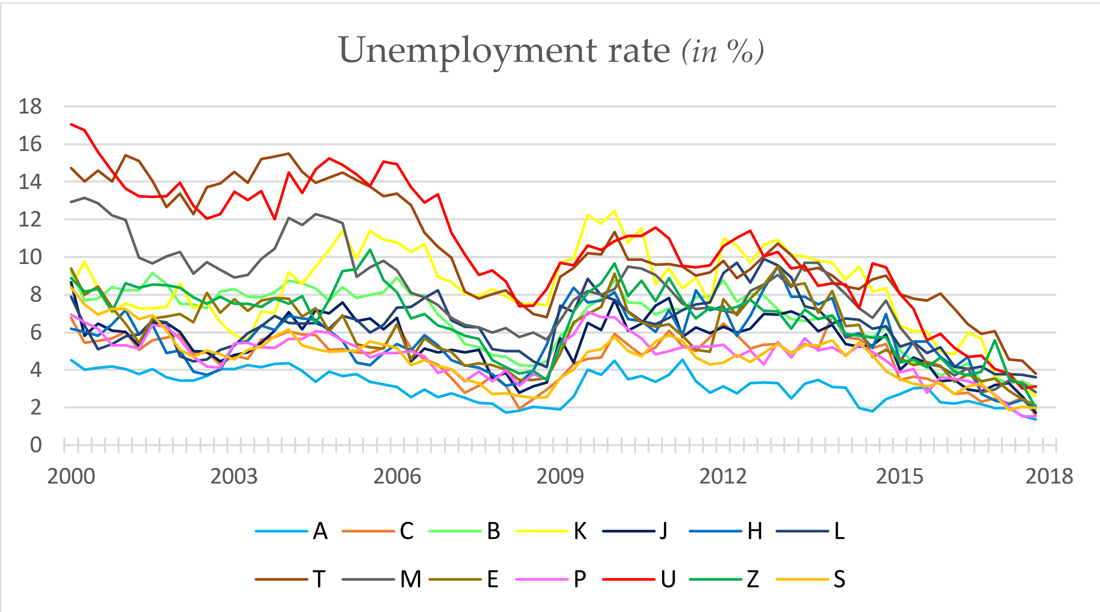
Source: CNB, IRI

Figure A.5: Construction cost in Czech regions



Source: CZSO

Figure A.6: Unemployment rate in Czech regions



Source: CZSO

Appendix B: Tables

Table B.1: Levin-Lin-Chu unit root test

Variable	*t*-statistic*	*p-value*
apartment price	-5.4180	0.0215**
building plot price	-5.1806	0***
new construction	-24.8718	0***
existing apartments	-2.5453	0.0291**
construction cost	-4.9752	0.0031***
wages	-1.7179	0.6115
unemployment	-3.8041	0.9938
marriages	-0.0011	0***
divorces	-11.8423	0***
natural growth of population	-18.5914	0***
net migration	-12.1229	0***
age structure	5.6228	1
reer	-9.6174	0***
pribor	-7.9662	0***
mortgages	-2.3081	0.7747
rent	1.1836	0.8817
infrastructure	-5.3359	0.0028***
urbanization	-5.1675	0***
log(apartment price)	-7.4291	0***
log(building plot price)	-7.0783	0***
log(wages)	-6.0130	0***
log(rent)	-4.0604	0.1714

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***

Table B.2: Breitung unit root test

Variable	*statistic*	*p-value*
apartment price	7.1583	1
building plot price	6.5930	1
new construction	-16.2060	0***
existing apartments	17.0823	1
construction cost	14.6710	1
wages	2.9366	0.9983
unemployment	1.0610	0.8557
marriages	-18.3392	0***
divorces	-14-0393	0***
natural growth of population	-7.2798	0***
net migration	-13.0171	0***
age structure	18.9813	1
reer	2.6030	0.9954
pribor	2.9155	0.9982
mortgages	-0.2180	0.4137
rent	4.8236	1
infrastructure	5.4041	1
urbanization	12.3542	1
log(apartment price)	6.6891	1
log(building plot price)	6.6359	1
log(wages)	2.6742	0.9963
log(rent)	4.6171	1

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***

Table B.3: Im-Pesaran-Shin unit root test

Variable	*statistic*	*p-value*
apartment price	-0.8613	0.9981
building plot price	-1.4195	0.6697
new construction	-7.1796	0***
existing apartments	-0.9986	0.9962
construction cost	-4.9570	0***
wages	-1.2866	0.8307
unemployment	-1-6452	0.2851
marriages	-8.3637	0***
divorces	-6.3774	0***
natural growth of population	-5.4816	0***
net migration	-5.0605	0***
age structure	5.7864	1
reer	-2.2037	0.0013***
pribor	-1.9594	0.024**
mortgages	-1.5565	0.4101
rent	-0.6821	0.9998
infrastructure	-0.3796	1
urbanization	-6.1475	0***
log(apartment price)	-1.8911	0.053*
log(building plot price)	-2.2061	0.0027***
log(wages)	-1.8625	0.0592*
log(rent)	-1.0521	0.9743

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***

Table B.4: Augmented Dickey-Fuller unit root test

Variable	*Choi ADF Test*		*Fisher ADF Test*	
	Z-statistic	*Prob.*	*statistic*	*p-value*
apartment price	0.1648	0.5655	22.0312	0.7798
building plot price	-0.1715	0.4319	25.5400	0.5983
new construction	-0.1618	0.4357	50.1214	0.0063***
existing apartments	3.0233	0.9987	27.5385	0.4891
construction cost	1.1910	0.8832	13.1657	0.9921
wages	8.6812	1	1.7788	1
unemployment	0.7846	0.7837	20.3329	0.8520
marriages	-0.8823	0.1888	23.8036	0.6919
divorces	-1.5497	0.0606*	91.0195	0***
natural growth of population	-2.1313	0.0165**	32.8564	0.2411
net migration	-7.5481	0***	169.9660	0***
age structure	7.5562	1	5.8264	1
reer	-3.9835	0***	60.2272	0.0004***
pribor	-2.4991	0.0062***	38.5828	0.0879*
mortgages	2.3504	0.9906	10.9466	0.9984
rent	2.9491	0.9984	7.3117	0.9999
infrastructure	-0.5552	0.2894	32.1895	0.2669
urbanization	-5.4641	0***	120.5900	0***
log(apartment price)	-2.1916	0.0142**	36.7304	0.1249
log(building plot price)	-2.6286	0.0043***	51.3565	0.0045***
log(wages)	4.5111	1	10.3959	0.9990
log(rent)	1.4387	0.9249	10.3730	0.9973

******* - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***

Table B.5: Phillips-Perron unit root test

Variable	*Choi PP Test*		*Fisher PP Test*	
	Z-statistic	*statistic*	*statistic*	*p-value*
apartment price	1.7318	0.9583	12.7101	0.9941
building plot price	-0.3263	0.3721	30.2508	0.3513
new construction	-18.6743	0***	421.0200	0***
existing apartments	2.3937	0.9917	34.8308	0.1748
construction cost	-4.9813	0***	66.9464	0***
wages	3.7649	0.9999	5.2637	1
unemployment	-0.4407	0.3297	26.8074	0.5288
marriages	-16.2942	0***	336.967	0***
divorces	-16.3112	0***	341.651	0***
natural growth of population	-15.5359	0***	317.656	0***
net migration	-13.7893	0***	268.758	0***
age structure	15.8338	1	0.00243	1
reer	-2.9911	0.0014***	43.4293	0.0316**
pribor	-1.7584	0.0393**	31.9751	0.2755
mortgages	1.7681	0.9615	11.7109	0.9971
rent	3.5049	0.9998	6.0965	1
infrastructure	3.1611	0.9992	10.3010	0.9991
urbanization	-6.2769	0***	141.842	0***
log(apartment price)	-1.1076	0.1340	29.6549	0.3799
log(building plot price)	-3.4434	0.0003***	70.1710	0***
log(wages)	-0.6868	0.2461	24.3207	0.6645
log(rent)	2.0459	0.9796	9.63609	0.9985

******* - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***

Table B.6: Hadri LM unit root test

Variable	*statistic*	*p-value*
apartment price	125.8186	0***
building plot prices	155.6082	0***
new construction	8.5403	0***
existing apartments	165.6621	0***
construction cost	148.4442	0***
wages	162.8432	0***
unemployment	79.5510	0***
marriages	-1.9048	0.9716
divorces	70.7971	0***
natural growth of population	49.9763	0***
net migration	11.6721	0***
age structure	130.9359	0***
reer	92.7186	0***
pribor	126.2097	0***
mortgages	130.4259	0***
rent	117.6000	0***
infrastructure	156.9156	0***
urbanization	154.5530	0***
log(apartment price)	121.5103	0***
log(building plot price)	153.6541	0***
log(wages)	159.6103	0***
log(rent)	126.2818	0***

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level*

Table B.7: Granger causality test for divorce rate

Null Hypothesis:	*Obs.*	*F-Stat.*	*Prob.*
divorce rate does not Granger cause apartment price	938	1.6304	0.1644
apartment price does not Granger cause divorce rate		3.0352	0.0168**

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level*

Table B.8: Panel dynamic OLS regression results - before and after crisis

log(apartment price)	*2000Q1-2008Q4*	*2009Q1-2017Q4*	*2000Q1-2017Q4*
log(building plot prices)	0.1756 (0.0797) **	0.4834 (0.0513) ***	0.3250 (0.0354) ***
log(wages)	1.3474 (0.2215) ***	0.6149 (0.0694) ***	0.6518 (0.0378) ***
unemployment	-0.0305 (0.0106) ***	-0.0014 (0.0081)	-0.0200 (0.0049) ***
marriages	-0.0235 (0.2171)	0.2427 (0.1814)	0.0180 (0.0255)
natural growth of population	-0.3830 (0.0817) ***	0.1063 (0.0562) *	-0.0488 (0.0376)
net migration	0.0124 (0.0155)	0.0184 (0.0185)	0.0430 (0.0076) ***
age structure	-0.0818 (0.0249) ***	-0.0202 (0.0086) **	-0.0090 (0.0041)
REER	0.0122 (0.0039) ***	0.0106 (0.0022) ***	0.0099 (0.0014) ***
ECT	-0.11 (0.04) ***	-0.05 (0.03)	-0.03 (0.01) ***
adjusted R^2	0.92	0.97	0.85

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level***
Standard errors in parentheses

Table B.9: Panel dynamic OLS regression results by time period

log(apartment price)	*2000Q1 - 2002Q4*	*2003Q1 - 2004Q4*	*2005Q1 - 2006Q4*	*2007Q1 - 2008Q4*	*2009Q1 - 2014Q4*	*2015Q1 - 2017Q4*	*2000Q1 - 2017Q4*
log(building plot prices)	0.3318 (0.0815) ***	0.2427 (0.0404) ***	0.2035 (0.0531) ***	0.1811 (0.0539) ***	0.4268 (0.0323) ***	0.5711 (0.0849) ***	0.3250 (0.0354) ***
log(wages)	2.3088 (0.3280) ***	1.7393 (0.1287) ***	0.8890 (0.1817) ***	1.0076 (0.1348) ***	0.6672 (0.0535) ***	0.9123 (0.1685) ***	0.6518 (0.0378) ***
unemployment	0.0079 (0.0113)	-0.0135 (0.0068) **	-0.0222 (0.0093) **	-0.0560 (0.0096) ***	-0.0209 (0.0042) ***	0.0035 (0.0160)	-0.0200 (0.0049) ***
marriages	0.1139 (0.0634) *	0.0016 (0.0303)	0.0183 (0.0277)	0.0297 (0.0157) *	-0.0496 (0.0127) ***	0.0480 (0.0177) ***	0.0180 (0.0255)
natural growth of population	0.0605 (0.1496)	-0.0673 (0.0585)	-0.0912 (0.0669)	-0.0420 (0.0611)	0.1294 (0.0264) ***	-0.0792 (0.0615)	-0.0488 (0.0376)
net migration	0.0936 (0.0590)	0.0369 (0.0135) ***	0.0342 (0.0115) ***	-0.0082 (0.0053)	0.0323 (0.0083) ***	0.0552 (0.0251) **	0.0430 (0.0076) ***
age structure	-0.2174 (0.0417) ***	-0.0518 (0.0135) ***	-0.0111 (0.0220)	-0.0180 (0.0151)	-0.0123 (0.0086)	-0.0552 (0.0197) ***	-0.0090 (0.0041)
REER	0.0001 (0.0030)	-0.0565 (0.0062) ***	0.0025 (0.0075)	0.0015 (0.0016)	0.0082 (0.0024) ***	-0.0012 (0.0050)	0.0099 (0.0014) ***
adjusted R^2	0.87	0.79	0.72	0.71	0.93	0.98	0.85

****** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level*

Standard errors in parentheses

IES Working Paper Series

2020

1. Tomas Kucera: *Cognitive Bias Mitigation: How to Make Decision-Making Rational?*
2. Tomas Kucera: *Are Employment Effects of Minimum Wage the Same Across the EU? A Meta-Regression Analysis*
3. Petr Hanzlik, Petr Teply: *Institutional and Other Determinants of the Net Interest Margin of US and European Banks in a Low Interest Rate Environment*
4. Michal Hlavacek, Ilgar Ismayilov, Ayaz Zeynalov: *Reassessment of the Fiscal Multiplier in Developing Countries: Regime-Switching Model*
5. Evzen Kocenda, Karen Poghosyan: *Nowcasting Real GDP Growth: Comparison between Old and New EU Countries*
6. Diana Zigraiova, Tomas Havranek, Jiri Novak: *How Puzzling Is the Forward Premium Puzzle? A Meta-Analysis*
7. Barbora Malinska: *Time-Varying Pricing of Risk in Sovereign Bond Futures Returns*
8. Shahriyar Aliyev, Evzen Kocenda: *ECB Monetary Policy and Commodity Prices*
9. Roman Kalabiska, Michal Hlavacek: *Regional Determinants of Housing Prices in the Czech Republic*

All papers can be downloaded at: http://ies.fsv.cuni.cz.

Univerzita Karlova v Praze, Fakulta sociálních věd
Institut ekonomických studií [UK FSV – IES] Praha 1, Opletalova 26
E-mail : ies@fsv.cuni.cz http://ies.fsv.cuni.cz