JEL Classification: O18, R11, R13, R21, R31
Keywords: apartment prices, regional analysis, residential real estate, panel regression

Regional Determinants of Housing Prices in the Czech Republic

Roman KALABISKA - Institute of Economic Studies, Faculty of Social Sciences, Charles University & Office of the Czech Fiscal Council, Czech Republic (r.kalabiska@seznam.cz) *corresponding author*

Michal HLAVACEK - Institute of Economic Studies, Faculty of Social Sciences, Charles University & Office of the Czech Fiscal Council, Czech Republic

Abstract

This paper examines the behaviour of housing prices and identifies their determinants across Czech regions from 2000 to 2019. The effect of a wide range of variables on apartment prices is analysed 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 quantify the speed of price adjustment in the short run. In order to check the robustness of the joint model, several regions with unique characteristics are excluded from the sample and analysed separately. Our results suggest that apartment prices are driven mainly by wages, unemployment rate and migration. We also found a large positive effect of building plot prices in high-income regions, while labour force factors (wages and age structure) seem to be more critical in low-income regions. The results of the joint error correction model suggest that shocks out of equilibrium are absorbed after approximately two years.

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. One of the most critical factors that affect property price is location. That is why in this paper, a special attention is paid to the region in which a property is located. For all 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 a robustness check to show which role various factors play in different regions.

The purchase of housing is a major component of household expenditures. Therefore understanding the dynamics of real estate prices and their 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

https://doi.org/10.32065/CJEF.2022.01.01

The authors would like to thank the anonymous reviewers for their helpful suggestions and comments.
We would like to acknowledge support from the SVV 260 597 project of the Charles University.

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) emphasises that on a large scale, residential property financing and sharp price corrections can undermine financial stability and lead to a slowdown in economic activity. These interconnections became apparent after the financial crisis of 2007/2008. Due to low-interest rates and a complex securitisation 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 2011). 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 that determine the price of residential real estate. The analysis uses a panel dataset covering a wide range of variables across 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 minimises 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. Previous analysis of the real estate price determinants usually considers only aggregate data for the whole country, while our study is one of the first considering the regional dimension of the real estate prices in the Czech Republic. In comparison to the only previous comprehensive study of regional determinants of real estate prices so far Hlaváček and Komárek (2011), we use a longer quarterly set of data that enables us to use more advanced methods of analysis as well as capture endogeneity in our data.[1]

2. Housing Price Determinants

In equilibrium, property price can be generally explained by fundamental factors, which affect both demand and supply. The supply of housing depends primarily on the profitability of construction firms. Following Égert and Mihaljek (2007), the overall construction cost includes building plot prices, cost of material and wages of construction workers. The relationship between the prices of real estate

---

[1] Hlaváček and Komárek (2011) use yearly panel data for period 1998-2008.

and building plots is apparent. As Hlaváček and Komárek (2011) explain in their paper, this variable suffers from an endogeneity problem. Each rise in 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. As a proxy for construction costs, we will use the *apartment construction price index*, which is provided by the CZSO. Furthermore, we will include the apartment stock and the number of newly built apartments in order to capture the construction activity and overall supply of housing. Alternatively, the number of construction permits could be used instead of the number of new apartments. They are expected to be highly correlated since one is essentially a lagged series of the other. The time difference between them is the length of the administrative and construction process. However, it does not have any quantifiable time nor regional dimension. Therefore it will not be considered in this paper. Hlaváček and Komárek (2011) established that supply factors are not important in determining real estate prices. However, due to the length of legislation and construction processes, changes in supply factors may affect real estate prices with a long lag.

Since the demand for housing is dependent on the disposable income of households, labour market factors are expected to play a key role in determining housing prices. As an increase in monthly wages leads to the accumulation of wealth and increases the availability of a mortgage loan, a positive effect on property prices is expected. Above that, we expect stronger demand for housing and higher prices in regions with higher average wages. We will further consider the unemployment rate in our analysis as the primary indicator of labour market health. Keeping the average wage constant, higher unemployment rate decreases the aggregate disposable income, which leads to lower purchasing power of households and a 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 results of Hlaváček and Komárek (2011) 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 (2018), the unemployment rate turned out to be insignificant as well.

Demographic factors describe the composition of the population and affect housing prices either directly or indirectly through the labour market. Furthermore, they also help to determine what types of properties are in demand (Asal 2018). Hlaváček and Komárek (2011) established the significance of population growth. They used both *natural population growth and population growth 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. A positive relationship between population growth and real estate prices was confirmed by Dröes and van de Minne (2017), Capozza et al. (2002), etc.

Apart from the size of the population, qualitative factors, such as the age structure of the population or household formation play an important role in the real estate market. Regions with a higher share of productive age population are expected to exhibit higher property prices due to stronger demand. We can find evidence of this effect in Dröes and van de Minne (2017), Belke and Keil (2018), etc. In this paper, we will use data for the *share of population aged 15-64* in each region published by CZSO as a proxy for age structure. 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 an additional dwelling. This relationship was confirmed by Hlaváček and Komárek (2011). 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 marriages may influence the data strongly.

A somewhat less common factor examined in academic literature is international competitiveness. An improvement of a country's competitiveness due to lower inflation or a weak exchange rate relative to foreign countries can attract foreign investors and increase domestic property prices. Asal (2018) used the *real effective exchange rate* (REER) and found that gains in trade competitiveness lead to a significant appreciation of house prices. In Czech literature, Hlaváček and Komárek (2009) 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 REER in the analysis. Its expected impact on apartment prices is negative, as an increase in REER implies that exports became more expensive relative to imports. Thus, the country experiences a loss in trade competitiveness.

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 the number of mortgages on property prices has been established by Égert and Mihaljek (2007), Belke and Keil (2018), Asal (2018) and others. As for Czech literature, Čadil (2009), Hlaváček and Komárek (2011) 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 *volume 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.

High mortgage rates and low rental prices can incentivise householders to live in a rented property rather than their own. A 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 (2011) argue that real estate prices can also influence rents. The increase in housing prices lowers the availability of owning property while simultaneously increases demand for rental properties and thereby increases rents. Because of this, we expect rents to be endogenous.

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. ar Dröes and van de Minne (2017) use an error correction model to show that the time needed to absorb shocks out of equilibrium varied from 0 to 20 years, with an average of 3 years. This suggests 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.

3. 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 relatively small in economic literature. Moreover, because of the existing historical influence of planned economy and transition processes, not a lot of research of property price determinants has been done in CEE countries until recent years (Égert and Mihaljek 2007; Hlaváček and Komárek 2011; Cohen and Karpavičiūtė 2017). Such studies also use data with shorter time periods than those of developed countries, such as Capozza et al. (2002), Schnure (2005) or Xu and Tang (2014). For all these reasons, the results may differ substantially based on the dataset used in each study.

In the Czech Republic, authors mostly analyse the real estate market using aggregate data for the whole country, for instance Čadil (2009), or focus on one city, such as Reichel and Zimčík (2018). Although Zemčík (2011) uses annual data for major cities of the Czech Republic and monthly data for major districts of Prague, the focus of this paper is to analyse the relationship between *real estate prices* and *rents* and identify areas with overvalued apartments.

The first comprehensive study of regional determinants of real estate prices in the Czech Republic is Hlaváček and Komárek (2011). This study includes a panel regression analysis of apartment prices across the Czech regions, 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. The authors also decided to estimate two 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*. As for other demand factors, coefficients of the *unemployment rate*, *growth in market rent* and *growth of average monthly wage* were all 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 the exponential growth of housing loans in 2002-2008 irrespective of developments on the real estate market, but they expect the standard relationship to restore after their observed period. Due to such unexpected result, this relationship will be tested in our paper in a later time period. 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 effect 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 authors use annual data, the time series in their panel regression is 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). The authors used panel dynamic OLS with error correction model to study determinants of real estate prices in 19 developed OECD countries and eight 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 generally observed higher in transition economies than in developed countries, which suggests that adjustment of property prices to the equilibrium is faster in CEE countries.

Another recent study of real estate prices determinants was conducted by Belke and Keil (2018). 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 that 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, the authors recommend testing the causality of variables before running the regression.

4. Data and Methodology

4.1 Real Estate Prices in the Czech Republic

The core variable of interest in our analysis will be apartment price. 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 realised prices. Supply prices 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 2011)

The data on transfer prices in the Czech Republic are collected by CZSO and published in "Prices of Observed Types of Real Estate" (CZSO 2020). 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 is 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 the municipality and the degree of wear of a given type of real estate. The main disadvantage can be seen in the delay of approximately one year with which these data are published. Despite that, we decided to use *transfer prices* in this paper due to their transparent methodology and complete regional and quarterly coverage. More specifically, transfer prices of *apartments* will be used due to higher homogeneity and higher number of transactions of apartments than family houses.[2]

The plot of apartment prices in the Czech regions is available in Appendix, part 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. Since 2015, prices have been increasing rapidly after almost six years of stagnation. In the fourth quarter of 2019. the average transfer price of apartments in the Czech Republic reached 32,250 CZK per square meter. In comparison, at the end of 2015, the average price was 21,554 CZK per square meter. That means a 49.6% increase over the four-year period. An essential 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

[2] The detailed discussion about advantages of transfer prices over supply prices and advantages of apartment prices over house prices can be found in Kalabiška and Hlaváček (2020).

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.

4.2 Data Description

Similarly to transfer prices of apartments ("*apartment prices*"), most of the data on explanatory variables are available from publications of CZSO or its public database. Data on building plot prices (per square meter), population and demographic variables, data on the number of apartments, construction cost index, average monthly wages and unemployment rate are all collected from CZSO. Information on the volume of mortgages and their regional breakdown is provided by MMR. All above-mentioned variables cover the time period from 2000Q1 to 2019Q4 for all Czech regions. REER, interest rates and rents (per square meter) are monitored by CZSO. 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 because it has its county seat in Prague, which is a separate region. The time period also differs as the available data on rents end in 2018Q2. Above that, the usability of rents in a regional analysis may be considered doubtful, as the data only include information for one municipality in each region.

Table 1 Descriptive Statistics

Variable	*Source*	*Obs.*	*Mean*	*Standard Deviation*		
				Overall	*Between*	*Within*
Apartment price	*CZSO*	*1120*	*18868.77*	*10208.78*	*8515.47*	*6103.93*
Building plot price	*CZSO*	*1120*	*2220.62*	*1475.99*	*1336.36*	*724.49*
*New construction (per 1,000 inhabitants)**	*CZSO*	*1120*	*0.68*	*0.32*	*0.26*	*0.20*
Apartment stock (per 1,000 inhabitants)	*CZSO*	*1120*	*386.71*	*19.17*	*17.08*	*9.88*
Construction cost index	*CZSO, CNB*	*1120*	*86.30*	*13.48*	*10.98*	*8.40*
*Wages**	*CZSO*	*1120*	*20864.55*	*5653.51*	*2327.47*	*5219.57*
Unemployment	*CZSO*	*1120*	*6.13*	*3.05*	*2.05*	*2.34*
*Marriages (per 1,000 inhabitants)**	*CZSO*	*1120*	*1.21*	*0.26*	*0.04*	*0.26*
*Divorces (per 1,000 inhabitants)**	*CZSO*	*1120*	*0.69*	*0.13*	*0.07*	*0.11*
*Natural population growth (per 1,000 inhabitants)**	*CZSO*	*1120*	*-0.08*	*0.99*	*0.14*	*0.99*
Migration (per 1,000 inhabitants)	*CZSO*	*1120*	*0.46*	*1.15*	*0.79*	*0.87*
Age structure	*CZSO*	*1120*	*68.85*	*2.41*	*0.52*	*2.37*
REER	*CNB*	*1120*	*102.02*	*9.99*	*0*	*10.05*
Interest rate	*CNB*	*1120*	*2.04*	*1.52*	*0*	*1.53*
*Mortgages (mil. CZK)**	*CNB*	*1120*	*1985.43*	*2675.04*	*2165.46*	*1657.92*
Rent	*CNB, IRI*	*962*	*96.30*	*25.12*	*17.83*	*18.48*

Notes: * - variable was adjusted for seasonality

The complete list of variables used in this study and their descriptive statistics can be found in Table 1. The dataset covers a time period of 20 years (80 quarters) from the first quarter of 2000 to the fourth quarter of 2019.[3] Combined with the panel size of 14 regions, it makes a total of 1,120 observations for all variables, except for rents, as discussed above.

4.3 Granger Causality

Let us consider the causal relationships in our model. As real estate prices affect economic variables, we will apply the 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 Three. The Granger (1969) causality test is defined as

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

The null hypothesis for the test is that lagged values of X do not have explanatory power on variation in Y and *vice versa*. Since some of our variables are I(1)[4], we will use first differences in order to meet the stationarity assumption of the Granger causality test. The lag length used for the test is 14, based on the results of lag length selection tests provided in Appendix, part 2. The Granger causality test (see Appendix, part 3) revealed that there are mostly mutual causal relationships between apartment prices and independent variables. An interesting result is that apartment price Granger causes 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 the value of the property. On the other hand, one can assume that rent does not create enough pressure on apartment prices to affect them. Furthermore, the causal link from natural population growth towards apartment prices has not been found either. This may be expected as in empirical literature, migration seems to be the leading regional driver of housing price dynamics out of population factors. Because these inferences have solid logical foundations, both rent and natural population growth will be excluded from further analysis.

5. Empirical Framework

For describing the long-run relations between cointegrated[5] 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 a 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, the DOLS estimator is

[3] Contrary to Hlaváček and Komárek (2011) we do not use data for 1998 and 1999 due to the change of the publication' structure and complicated comparability of the data with later periods.
[4] The discussion of supplementary tests for stationarity can be found in Appendix, part 4.
[5] The discussion of supplementary tests for cointegration can be found in Appendix, part 5.

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 that 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:

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

where $Y_{i,t}$ is *log(apartment price)*, β's are estimated coefficient of explanatory variables X. The maximum lag and lead length will be determined by HQC as it is the most suitable for the given dataset (see Appendix, part 2). Together with PDOLS, an error correction model (ECM) is employed to capture the short-run dynamics of apartment prices. We apply the two-step ECM procedure as described by Engle and Granger (1987), which uses the error correction term (ECT) that will be estimated by PDOLS. 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.

5.1 Regression Results

Due to a relatively large number of variables that enter our model, the estimation will be made initially on a restricted benchmark model that includes demand-side fundamentals – *wages, unemployment* and *age structure*. These variables were chosen because the disposable income of households and the size of the labour force were established as important factors in determining prices of real estate in the 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 the relevant literature and has been applied by Égert and Mihaljek (2007), Huynh-Olesen et al. (2013), etc.[6]

Table 2 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.*

[6] In order to check robustness of chosen procedure, we also performed backward stepwise regression (see PDOLS(14) in Table 4).

Table 2 Panel Regression Results of the Baseline Model Specification

log(apartment price)	*Pooled OLS*	*Fixed Effects*	*Random Effects*	*PDOLS*	*FMOLS*	*ARDL*	*GMM (1)*	*GMM (2)*
log(apartment price) (-1)								0.9775 (0.0049) ***
log(wages)	1.98 (0.0401) ***	1.55 (0.0193) ***	1.53 (0.0193) ***	1.11 (0.0565) **	1.28 (0.0283) ***	1.69 (0.0511) ***	1.98 (0.0401) ***	-0.0132 (0.0095)
unemployment	-0.0721 (0.0031) ***	-0.0228 (0.0018) ***	-0.0231 (0.0018) ***	-0.0361 (0.0047) ***	-0.0740 (0.0176) ***	-0.0445 (0.0045) ***	-0.0721 (0.0031) ***	-0.0050 (0.0006) ***
age structure	0.0707 (0.0045) ***	0.0490 (0.0021) ***	0.0492 (0.0021) ***	0.0575 (0.0046) ***	0.0555 (0.0050) ***	0.0413 (0.0040) ***	0.0707 (0.0045) ***	0.0002 (0.0008)
constant	-7.4764 (0.6249) ***	-6.6416 (0.3047) ***	-6.6479 (0.3106) ***				-7.4764 (0.6249) ***	0.3819 (0.1135) ***
adjusted R2	0.72	0.96	0.90	0.96	0.95		0.72	0.99
overall R2	0.73	0.67	0.67				0.73	0.99
within R2		0.90	0.90					
between R2		0.60	0.60					
rho		0.88	0.80					

*Notes: *** - significant at 1% level, ** - significant at 5% level, * - significant at 10% level. Rho shows the fraction of variance caused by the individual term. Standard errors in parentheses.*

As we are dealing with dynamic panel data, which likely contain some form of endogeneity and cointegration, standard panel estimators are presumably biased. Nevertheless, the similarity of their results to those of dynamic estimators suggests that the endogeneity and cointegration do not affect the results too much. We can also notice that both fixed effects and random effects estimators yield very similar results as well. This could be explained by the Hausman test (see Table 3), which was unable to reject the null hypothesis, that the difference in coefficients is non-systematic. Under this hypothesis, both FE and RE estimators are consistent, but only RE is efficient and is therefore preferred to FE.

Table 3 Hausman Test

Null hypothesis	*Coef.*	*Prob.*
Cross-section random	*5.69*	*0.1277*

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

The robustness of our chosen estimator, PDOLS, is supported by similar results of alternative dynamic estimators – FMOLS, ARDL and GMM. Although ARDL is more flexible with variables of a different order of integration, dynamic OLS is more suitable and used more often in empirical literature for estimating long-run coefficients. We also include a second specification of the GMM estimator, which includes lagged values of apartment price as an explanatory variable. As its coefficient is close to 1, we can infer that the autocorrelation of the dependent variable is very strong and in this specification, coefficients of remaining variables are likely biased because of it.

In the baseline model estimated by PDOLS, the coefficient of wages indicates that a 1 % increase in the average monthly wage would cause apartment prices to increase, on average, by 1.45 %. 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 percentage point, the apartment prices rise by approximately 3.6 %. Both these coefficients 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, which is the proxy for age structure, increases by one percentage point, the apartment prices increase by 5.8 % due to higher demand.

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 4. Results of these regressions are denoted PDOLS(2) – PDOLS(11).

Firstly, supply factors are added in models (2) through (5). While the impact of apartment stock on apartment prices is significant at 1% level, the coefficients of other supply factors are insignificant. As the relationship between building plots and apartments as assets is substitutive, the results suggest that building plot prices do not affect apartment prices in the long run. Similarly, the apartment price does not seem to react at all to the number of newly build apartments, as housing prices are rather rigid. In case of the construction cost index, 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, the construction cost index measures the cost of new apartments, which are inherently more expensive than second-hand apartments. Therefore, these two variables are not easily comparable. On the other hand, the apartment stock has the correct negative sign and is statistically significant. Thus, it will be included in the extended model.

Models (6) through (8) examine the effects of demographic factors on apartment prices. While the results suggest that demand for housing is not affected by marriages, the coefficients of divorces and migration are both positive and significant at 1% level. As these results support our expectations, both variables will be included in the extended regression 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 in the apartment price. Although the effect is relatively small, the sign of the coefficient goes against economic theory, as discussed in Chapter Two. On the other hand, the coefficient of interest rate is not significant. Similarly, model (11) suggests that apartment prices are not affected by mortgages. Therefore, both variables will not be included in the extended regression model.

Table 4 Panel Regression Results

log(apartment price)	PDOLS (1)	PDOLS (2)	PDOLS (3)	PDOLS (4)	PDOLS (5)	PDOLS (6)	PDOLS (7)	PDOLS (8)	PDOLS (9)	PDOLS (10)	PDOLS (11)	PDOLS (12)	PDOLS (13)	PDOLS (14)
log(building plot price)		-0.0181 (0.0894)												
new construction			0.0003 (0.0502)											
apartment stock				-0.0070 (0.0020) ***								-0.0034 (0.0019) *	-0.0033 (0.0017) **	-0.0057 (0.0016) ***
construction cost index					-0.0009 (0.0048)									
log(wages)	1.4511 (0.0565) ***	1.4235 (0.1176) ***	1.4407 (0.0618) ***	1.6017 (0.0691) ***	1.4265 (0.1848) ***	1.4684 (0.0658) ***	1.4649 (0.0503) ***	1.3807 (0.0563) ***	1.1026 (0.1214) ***	1.4753 (0.0656) ***	1.3670 (0.0986) ***	1.1670 (0.1370) ***	0.9615 (0.0924) ***	0.8203 (0.0973) ***
unemployment	-0.0361 (0.0047) ***	-0.0373 (0.0052) ***	-0.0360 (0.0049) ***	-0.0322 (0.0044) ***	-0.0327 (0.0052) ***	-0.0339 (0.0053) ***	-0.0350 (0.0042) ***	-0.0273 (0.0052) ***	-0.0328 (0.0047) ***	-0.0342 (0.0052) ***	-0.0357 (0.0047) ***	-0.0222 (0.0045) ***	-0.0244 (0.0039) ***	-0.0294 (0.0037) ***
marriages						0.0344 (0.0465)								
divorces							0.4448 (0.1048) ***					0.3404 (0.0920) ***	0.4491 (0.0864) ***	0.3740 (0.0866) ***
migration								0.0363 (0.0110) ***				0.0306 (0.0105) ***	0.0400 (0.0111) ***	0.0265 (0.0105) **
age structure	0.0575 (0.0046) ***	0.0567 (0.0052) ***	0.0565 (0.0058) ***	0.0454 (0.0049) ***	0.0561 (0.0066) ***	0.0566 (0.0046) ***	0.0479 (0.0050) ***	0.0453 (0.0051) ***	0.0325 (0.0090) **	0.0569 (0.0045) ***	0.0555 (0.0053) ***	0.0115 (0.0086)		
REER									0.0062 (0.0020) ***			0.0063 (0.0020) ***	0.0079 (0.0012) ***	0.0072 (0.0012) ***
interest rate										0.0060 (0.0069)				
log(mortgages)											0.0195 (0.0197)			0.0661 (0.0168) ***
ECT	-0.09 (0.01) ***	-0.08 (0.01) ***	-0.10 (0.01) ***	-0.10 (0.01) ***	-0.10 (0.02) ***	-0.10 (0.02) ***	-0.09 (0.02) ***	-0.11 (0.01) ***	-0.10 (0.01) ***	-0.09 (0.01) ***	-0.09 (0.01) ***	-0.12 (0.02) ***	-0.11 (0.01) ***	-0.13 (0.01) ***
R^2	0.96	0.96	0.96	0.97	0.97	0.96	0.97	0.97	0.96	0.96	0.96	0.97	0.97	0.97
adjusted R^2	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.97	0.96	0.97

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

In order to capture the short-term dynamics of apartment prices, an error correction term (ECT) was calculated based on the result of each regression. Each ECT was estimated using lagged residuals obtained from given PDOLS regression. As we can see in models (1) through (11), 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 -0.08 and -0.11, which suggests that in every time period, i.e. every quarter, there is between 8% and 11% adjustment towards the equilibrium.

Column (12) shows the results of the extended model regression estimated by PDOLS. The following variables were included based on the model selection procedure described above – apartment stock, wages, unemployment, divorces, migration, age structure and REER. 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. The coefficient of apartment stock is now significant at only 10% level, and age structure has lost its statistical significance altogether. All other variables are still significant at 1% level and have the same sign as in their respective reduced models. In order to correct for the uncertainty, we will re-estimate the model with age structure excluded. From the results in column (13) we can see that the coefficients changed marginally, and standard errors of most variables have decreased. Therefore, this model seems to be more robust.

Since this model was specified using forward stepwise selection, we also add a model (14), which was specified by backward stepwise elimination, in which a variable is eliminated in each step of the selection process. One can see that the results of both extended models do not differ much, with one notable exception[8]. Mortgages were estimated to have a positive effect on apartment prices. With each 1% increase in the volume of mortgage loans, apartment prices grow by 0.06 %. Furthermore, we can conclude that demand-side factors (wages, unemployment and migration) are important determinants of apartment prices. The divorce rate proved to have a positive effect on housing demand as well. Out of supply-side factors, only apartment stock proved to be important. Finally, the positive impact of REER on apartment prices has been confirmed in the extended regression. This analysis has also shown that other factors are of low importance in determining apartment prices.

5.2 Robustness Checks

In order to be sure about our results, additional regressions are performer on selected subsamples. The extended panel regression model PDOLS(14) from Section 5.1 is used to estimate models for the Czech Republic with outlier regions eliminated. From the original dataset, which includes data for all 14 Czech regions, we will exclude two groups of regions and estimate our model on each group separately using the same model selection procedure from Section 5.1. This will enable us to check the robustness of the extended model as well as examine different effects of housing price determinants in different regions.

[8] This can be caused by the omitted variable bias in forward stepwise regression.

Prague (A) has been identified as a "*high-income*" outlier for several reasons. Firstly, several 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 the 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 capital city. For that reason, this analysis would be best done on a district level rather than a regional 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 that differ from the rest are Moravskoslezský (T) and Ústecký (U) regions. In the 20th century, both were highly industrialised regions. Local coal mines supplied thousands of jobs, and large apartment buildings were provided for miners. However, in the 1990s the mining industry declined dramatically as coal mining was no longer profitable after the transformation of the 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 2000Q1-2019Q4, 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 the regional analysis are presented in Table 5, which is divided into three parts. The first three columns show results for models, where Prague is analysed separately [9] in column (1) by dynamic OLS, the model for Moravskoslezský and Ústecký regions is estimated by panel dynamic OLS in column (2) and the rest of Czech Republic (11 regions) is also analysed by panel dynamic OLS in column (3). Columns (4) through (6) show results for models where Středočeský region is excluded from CZ and is estimated together with Prague in column (4). For better readability of the table, the results for low-income regions are repeated in column (5).[10] In order to provide a direct comparison, the results of joint panel regression for all 14 regions conducted in Section 5.1 is provided in column (7).

[9] Due to a low amount of observations in the regression for Prague, we provide alternative estimations of this model in Appendix, part 6 as a robustness check.

[10] Note that columns (2) and (5) are repetitive. This is due to illustration purposes, since columns (1) – (3) together include all 14 regions and also columns (4) – (6) together include all 14 regions.

Table 5 PDOLS 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 price)	*0.3852* *(0.1580)* ***			*0.3698* *(0.1404)* ***			
apartment stock		*0.0269* *(0.0060)* ****	*-0.0054* *(0.0018)* ****		*0.0269* *(0.0060)* ****	*-0.0076* *(0.0021)* ****	*-0.0057* *(0.0016)* ****
log(wages)		*0.9135* *(0.1799)* ****	*0.8266* *(0.1071)* ****		*0.9135* *(0.1799)* ****	*0.9376* *(0.1145)* ****	*0.8203* *(0.0973)* ****
unemployment	*-0.0513* *(0.0234)* ***	*-0.0334* *(0.0106)* ****	*-0.0283* *(0.0044)* ****	*-0.0496* *(0.0176)* ****	*-0.0334* *(0.0106)* ****	*-0.0296* *(0.0042)* ****	*-0.0294* *(0.0037)* ****
divorces			*0.3468* *(0.0928)* ****			*0.3572* *(0.0888)* ****	*0.3740* *(0.0866)* ****
migration	*0.0116* *(0.0081)*	*0.0322* *(0.0251)*	*0.0338* *(0.0127)* ****	*0.0289* *(0.0143)* ***	*0.0322* *(0.0251)*	*0.0255* *(0.0133)* **	*0.0265* *(0.0105)* ***
age structure	*-0.0280* *(0.0102)* ****	*0.1400* *(0.0195)* ****		*-0.0262* *(0.0122)* ***	*0.1400* *(0.0195)* ****		
REER	*-0.0098* *(0.0037)* ****		*0.0070* *(0.0013)* ****	*0.0105* *(0.0037)* ****		*0.0065* *(0.0012)* ****	*0.0072* *(0.0012)* ****
interest rate		*0.0378* *(0.0144)* ****			*0.0378* *(0.0144)* ****		
log(mortgage)			*0.0639* *(0.0180)* ****			*0.0634* *(0.0169)* ****	*0.0661* *(0.0168)* ****
ECT	*-0.16* *(0.05)* ****	*-0.16* *(0.05)* ****	*-0.17* *(0.02)* ****	*-0.15* *(0.05)* ****	*-0.16* *(0.05)* ****	*-0.17* *(0.02)* ****	*-0.13* *(0.01)* ****
R^2	*0.95*	*0.97*	*0.96*	*0.99*	*0.97*	*0.95*	*0.97*
adjusted R^2	*0.95*	*0.96*	*0.95*	*0.98*	*0.96*	*0.95*	*0.97*

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

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), similarly to the rest of CZ, it was insignificant in the variable selection process and excluded from the final model. Due to the increasing shortage of buildable land in Prague, its availability together with building plot prices are expected to have a significant impact on apartment prices. On the other hand, there have been thousands of disused apartments available in Moravskoslezský and Ústecký regions in the early 2000s. 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 a substitutive market. For that reason, apartment stock seems to have a positive effect on apartment prices in low-income regions. Similar results are reported in models (4) through (6), which confirms our hypothesis.

The results suggest that wages do not have any significant impact on apartment prices in Prague and Středočeský region as opposed to 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, has a negative coefficient in these regions. On the other hand, the effects of the unemployment rate are quite similar throughout all estimated models and only differ for high-income regions.

Demographic factors seem to have rather uneven effects on apartment prices in each group of regions. Although the divorce rate does have a positive coefficient in the joint model as well as its derivations with outlier regions excluded, it is statistically insignificant in models for high-income and low-income regions. A similar inference can be made about migration. However, it is also significant in model (4). This could be explained by the fact that Středočeský region works as a hub because housing in Prague is often too high for reach, and commuting from Středočeský region to Prague is affordable.

We can also see that results for REER vary for high-income regions, as both experience demand from different types of investors. Whereas Prague attracts mainly foreign investors, Středočeský region is in this sense more similar to other Czech regions and attracts mostly domestic investors. Last but not least, one can notice rather mixed effects of the interest rate and mortgages on apartment prices. While the significance of mortgages was found in models for the rest of CZ as well as the joint model, interest rate seems to affect apartment prices only in low-income regions.

The larger size of the error correction term in regional models suggests that housing price can recover faster after shocks in individual regions. This result also makes sense from a regional policy view. The speed of recovery in all regional models lies between 15 % and 17 % of adjustment towards equilibrium price each quarter. At this rate, a short-term shock would be absorbed within two years.

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 a residual-based approach to describe price misalignments on three groups of regions – Prague, Moravskoslezský and Ústecký region and the rest of the country (11 regions). Although this analysis presents valuable insights on trends and stability of housing prices, there is no information provided about the different impacts of housing price determinants across regions.

6. Conclusions

Due to the complexity of the real estate market, housing price developments are influenced by many variables. Several methods were used in order to analyse 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 PDOLS, FMOLS, ARDL and GMM. Based on the nature of the available data, an extended PDOLS model was identified in order to describe the long-term equilibrium of the relationship between apartment prices and their determinants. Furthermore, we employed an error correction model to capture the short-run dynamics of apartment prices.

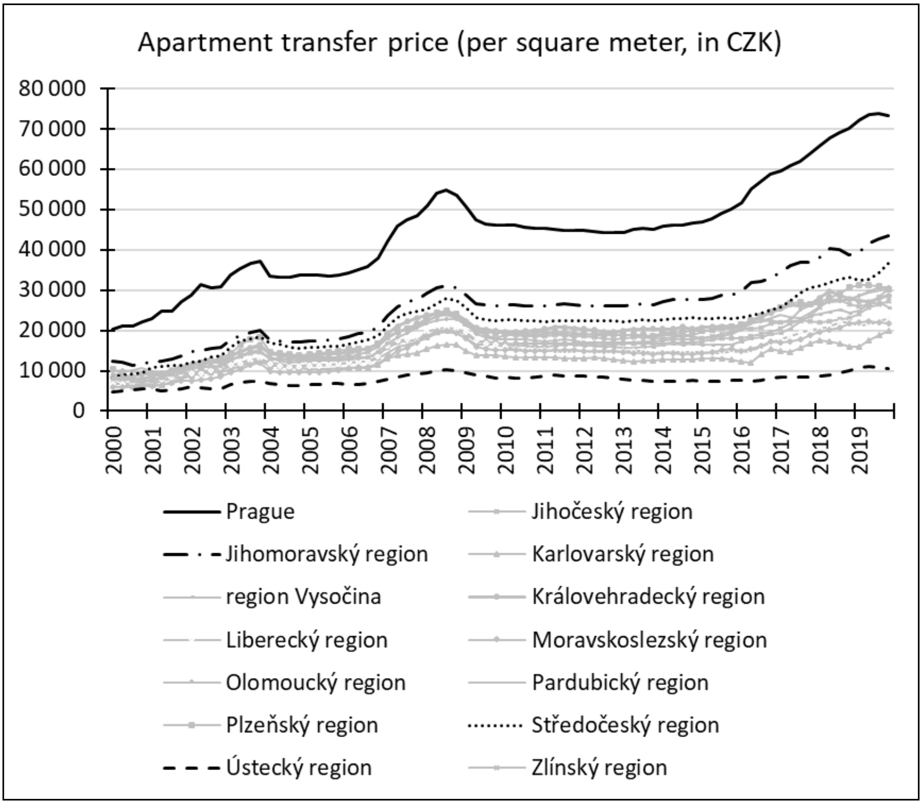
The PDOLS regression revealed that apartment prices in the Czech Republic are driven mainly by demand-side fundamentals: wages and unemployment. These results are consistent with empirical literature related to this paper, namely Hlaváček and Komárek (2011) and Čadil (2009). Of the supply factors, only apartment stock turned out to be significant. This finding adds valuable information to the existing literature, as it was supported by several robustness checks. Although the effects of demographic factors were rather mixed, positive effects of divorce rate and migration on apartment prices were established. The regression also showed that there is an unexpected positive relationship between REER and apartment prices. A positive effect of mortgages was also established in the joint model. 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 is not Granger caused by rents and natural population growth.

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 positive in Prague and Středočeský regions, while it was insignificant in others. On the other hand, labour force factors (wages and age structure) seem to be more important in low-income regions than they are in Prague. Furthermore, the regression established that the results for Prague are very similar to the results for both Prague and Středočeský regions when estimated together.

To the best of the authors' knowledge, this is the first application of panel data regression of real estate prices by region in the Czech Republic. The presented regressions could be further enhanced by analysing 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.

APPENDIX

1. Apartment Prices in Czech Regions



Source: Authors' calculations based on CZSO's data

2. Lag Length Selection

In order to be able to choose the correct number of AR lags, a vector autoregression (VAR) model had to be estimated. Some of the most common selection criteria, which have been used in relevant literature are the *Akaike information criterion* (AIC), *Schwarz information criterion* (SIC) and *Hannah-Quinn information criterion* (HQC). Liew (2004) established that with a relatively large sample (120 or more observations), HQC outperforms other criteria. In contrast, AIC is a better choice in a small sample (60 observations and less). As the dataset used in this paper includes 1,120 observations, it can be inferred that HQC is the most suitable. Nevertheless, all aforementioned information criteria selected the optimal lag length of 14.

Table A1 VAR Lag Order Selection Criteria

Lags	*AIC*	*SIC*	*HQC*
0	*0.1276*	*0.1501*	*0.1362*
1	*-3.9342*	*-3.9060*	*-3.9234*
2	*-4.4003*	*-4.3665*	*-4.3874*
3	*-4.4786*	*-4.4391*	*-4.4635*
4	*-4.4932*	*-4.4481*	*-4.4759*
5	*-4.4917*	*-4.4410*	*-4.4723*
6	*-4.4893*	*-4.4330*	*-4.4677*
7	*-4.4923*	*-4.4303*	*-4.4685*
8	*-4.5037*	*-4.4360*	*-4.4777*
9	*-4.5211*	*-4.4478*	*-4.4930*
10	*-4.5249*	*-4.4460*	*-4.4946*
11	*-4.5232*	*-4.4387*	*-4.4908*
12	*-4.5359*	*-4.4457*	*-4.5013*
13	*-4.5387*	*-4.4429*	*-4.5020*
14	*-4.5498**	*-4.4484**	*-4.5110**
15	*-4.5495*	*-4.4425*	*-4.5085*
16	-4.5488	-4.4361	-4.5056

*Notes: * - lag order selected by the criterion*

3. Granger Causality Tests

Null Hypothesis:	*F-Stat.*	*Prob.*
building plot price does not Granger cause apartment price	*2.7788*	*0.0005****
apartment price does not Granger cause building plot price	*3.4248*	*0.0000****
new construction does not Granger cause apartment price	*2.3417*	*0.0035****
apartment price does not Granger cause new construction	*6.4242*	*0.0000****
apartment stock does not Granger cause apartment price	*2.5678*	*0.0013****
apartment price does not Granger cause apartment stock	*5.2381*	*0.0000****
construction cost index does not Granger cause apartment price	*9.2514*	*0.0000****
apartment price does not Granger cause construction cost index	*17.1984*	*0.0000****
wages does not Granger cause apartment price	*7.8814*	*0.0000****
apartment price does not Granger cause wages	*8.0044*	*0.0000****
unemployment does not Granger cause apartment price	*2.3015*	*0.0042****
apartment price does not Granger cause unemployment	*5.1442*	*0.0000****
marriages does not Granger cause apartment price	*8.4282*	*0.0000****
apartment price does not Granger cause marriages	*5.9319*	*0.0000****
divorces does not Granger cause apartment price	*2.0163*	*0.0143***
apartment price does not Granger cause divorces	*2.4558*	*0.0021****
natural population growth does not Granger cause apartment price	*0.6795*	*0.7959*
apartment price does not Granger cause natural population growth	*1.8896*	*0.0241***
migration does not Granger cause apartment price	*8.7018*	*0.0000****
apartment price does not Granger cause migration	*7.0604*	*0.0000****
age structure does not Granger cause apartment price	*4.6300*	*0.0000****
apartment price does not Granger cause age structure	*5.0209*	*0.0000****
REER does not Granger cause apartment price	*13.3284*	*0.0000****
apartment price does not Granger cause REER	*15.5522*	*0.0000****
interest rate does not Granger cause apartment price	*14.0765*	*0.0000****
apartment price does not Granger cause interest rate	*11.2515*	*0.0000****
mortgages does not Granger cause apartment price	*6.9036*	*0.0000****
apartment price does not Granger cause mortgages	*9.2880*	*0.0000****
rent does not Granger cause apartment price	*1.5969*	*0.0743**
apartment price does not Granger cause rent	*4.7411*	*0.0000****

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

4. Stationarity Tests

This section presents unit root tests for checking the stability of the data. 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. 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 the first difference of the data. Another possibility is to use a logarithmic transformation.

Two of the most common tests are the Augmented Dickey-Fuller (1979, ADF) test and Phillips-Perron (1988, PP) test (see Tables A2 and A3). Both tests are available with two specifications. First is the standard Fisher-type test, which follows asymptotic Chi-square distribution. The 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 the null hypothesis of individual unit root process. The lag lengths for all tests were chosen by HQC.[11] 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 standardised t-bar statistic as $N \to \infty$, according to Monte Carlo simulations performed by Im et al. (2003), the IPS test performs better than other unit root tests in small samples. Among other tests available, Levin–Lin–Chu (2002, LLC) and Breitung (2000) both assume a common unit root process and rely on the assumption of cross-sectional independence. Therefore, these tests do not seem to be appropriate for our dataset (see Appendix, part 7).

Based on the results of performed unit root tests, we can conclude that *apartment price*, *construction cost index*, *unemployment*, *age structure*, *REER*, *interest rate* and *rent* are non-stationary, as this was reported by all three tests listed above. The results of other variables, with the exception of *marriages*, seem to be conclusive as well since at least two tests rejected the null hypothesis of a unit root presence at 5% significance level. As for *marriages*, while PP test rejected the null hypothesis of a unit root, the ADF and the IPS tests did not. Due to these mixed results, the following analysis must be conducted with caution.

[11] The detailed discussion about choosing the optimal lag length can be found in Appendix, part 2.

Table A2 Augmented Dickey-Fuller Unit Root Test

Variable	*Choi ADF Test*		*Fisher ADF Test*	
	Z-statistic	***Prob.***	***statistic***	***p-value***
apartment price	*-0.4496*	*0.3265*	*24.5017*	*0.6548*
building plot price	*-1.8054*	*0.0355***	*41.1489*	*0.0520**
new construction	*-7.8076*	*0.0000****	*171.3020*	*0.0000****
apartment stock	*-1.6611*	*0.0483***	*44.4288*	*0.0252***
construction cost index	*1.8584*	*0.9684*	*10.3778*	*0.9990*
wages	*-3.2282*	*0.0006****	*51.6972*	*0.0042****
unemployment	*0.3159*	*0.6240*	*22.9302*	*0.7365*
marriages	*5.4104*	*1.0000*	*2.9393*	*1.0000*
divorces	*-19.0771*	*0.0000****	*438.5300*	*0.0000****
natural population growth	*-4.8358*	*0.0000****	*75.3354*	*0.0000****
migration	*-8.1425*	*0.0000****	*143.0840*	*0.0000****
age structure	*-1.2042*	*0.1143*	*35.1189*	*0.1664*
REER	*-1.1939*	*0.1163*	*27.4755*	*0.4925*
interest rate	*-0.5114*	*0.3045*	*22.6306*	*0.7513*
mortgages	*-2.2427*	*0.0125***	*37.2007*	*0.1145*
rent	*2.6522*	*0.9960*	*12.3903*	*0.9888*
log(apartment price)	*-2.1363*	*0.0163***	*36.6217*	*0.1274*
log(building plot price)	*-2.7514*	*0.0030****	*52.3596*	*0.0035****
log(wages)	*0.2090*	*0.5828*	*18.9180*	*0.9006*
log(mortgage)	*-1.6697*	*0.0475***	*32.3174*	*0.2618*

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

Table A3 Phillips-Perron Unit Root Test

Variable	*Choi ADF Test*		*Fisher ADF Test*	
	Z-statistic	***Prob.***	***statistic***	***p-value***
apartment price	*2.6999*	*0.9965*	*8.3342*	*0.9999*
building plot price	*0.9470*	*0.8282*	*17.8005*	*0.9311*
new construction	*-13.8557*	*0.0000****	*265.6610*	*0.0000****
apartment stock	*1.6892*	*0.9544*	*15.5208*	*0.9724*
construction cost index	*2.4662*	*0.9932*	*8.2390*	*0.9999*
wages	*-5.5535*	*0.0000****	*77.0827*	*0.0000****
unemployment	*-0.4974*	*0.3094*	*24.5309*	*0.6532*
marriages	*-19.3037*	*0.0000****	*446.9840*	*0.0000****
divorces	*-19.6161*	*0.0000****	*458.0910*	*0.0000****
natural population growth	*-17.6254*	*0.0000****	*387.4060*	*0.0000****
migration	*-14.7633*	*0.0000****	*301.2610*	*0.0000****
age structure	*0.5382*	*0.7048*	*21.3534*	*0.8103*
REER	*1.5016*	*0.9334*	*11.8086*	*0.9968*
interest rate	*3.5440*	*0.9998*	*5.2772*	*1.0000*
mortgages	*-9.6244*	*0.0000****	*149.5350*	*0.0000****
rent	*2.9800*	*0.9986*	*9.4013*	*0.9988*
log(apartment price)	*0.2996*	*0.6178*	*18.2034*	*0.9209*
log(building plot price)	*-1.1513*	*0.1248*	*30.6684*	*0.3320*
log(wages)	*-3.7068*	*0.0001****	*52.1700*	*0.0037****
log(mortgage)	*-2.4308*	*0.0075****	*38.5143*	*0.0891**

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

Table A4 Im-Pesaran-Shin Unit Root Test

Variable	*W-statistic*	*p-value*
apartment price	*-0.4399*	*0.3300*
building plot price	*-1.9231*	*0.0272*****
new construction	*-9.3146*	*0.0000******
apartment stock	*-1.9632*	*0.0248*****
construction cost	*1.7393*	*0.9590*
wages	*-3.4888*	*0.0002******
unemployment	*0.1114*	*0.5443*
marriages	*5.0296*	*1.0000*
divorces	*-26.8410*	*0.0000******
natural population growth	*-4.8122*	*0.0000******
migration	*-8.4291*	*0.0000******
age structure	*-1.3404*	*0.0901****
REER	*-1.0774*	*0.1407*
interest rate	*-0.5637*	*0.2865*
mortgages	*-2.3374*	*0.0097******
rent	*2.5495*	*0.9946*
log(apartment price)	*-2.0329*	*0.0210*****
log(building plot price)	*-2.8683*	*0.0021******
log(wages)	*-0.1798*	*0.4287*
log(mortgage)	*-1.7824*	*0.0373*****

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

5. Cointegration Tests

After establishing that some of our variables are non-stationary, the next step is to find out whether or not there are any cointegrating relationships among them. A common approach in the 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) emphasises 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 the 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 by AIC, SIC and HQC.

Kao test was performed on two sets of I(1) variables (see Table A5). The full set includes *log(apartment price)*, *construction cost index*, *unemployment*, *marriages*, *age structure*, *REER* and *interest rate*. However, because one can make the argument that *marriages* could be stationary, we decided to create a restricted set with *marriages* excluded. Furthermore, rent was excluded from further analysis due to the results of the Granger causality test in Appendix 3. For both sets of variables, Kao test rejected the null hypothesis of no cointegration at 1% confidence level.

Table A5 Kao Cointegration Test

ADF	Full set		Restricted set	
	t-statistic	prob.	t-statistic	prob.
Schwarz IC	-8.8657	0.0000***	-8.8281	0.0000***
Hannah-Quinn IC	-3.1093	0.0009***	-2.9158	0.0018***
Akaike IC	-3.1093	0.0009***	-2.9158	0.0018***

*Notes: *** - 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 assumes common AR process and has the alternative hypothesis that all panels are cointegrated, whereas group statistic, which assumes individual AR uses the heterogeneous alternative of some cointegrating panels. Group statistics are the group mean extension of the panel version. Table A6 revealed that Pedroni test rejects the null hypothesis in both 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 A6 Pedroni Cointegration Test

Alternative hypothesis	Full set		Restricted set	
	ADF-Statistic	Prob.	ADF-Statistic	Prob.
Common AR coefs.	-4.7026	0.0000***	-5.2185	0.0000***
Individual AR coefs.	-5.8273	0.0000***	-6.4781	0.0000***

*Notes: *** - 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. Table A7 presents results of the Johansen cointegration test, which suggests that there are at least three cointegrating relationships among variables in the full set, as the null hypothesis of at most two cointegrating relationships was rejected at 1% significance level. In the restricted set, the hypothesis of at most one cointegrations was rejected. Therefore, we can conclude that in the restricted set there are at least two cointegrating relationships.

Table A7 Johansen Panel Cointegration Test

Hypothesis of no. of cointeg.	Full set		Restricted set	
	F-statistic	Prob.	F-statistic	Prob.
0	352.80	0.0000***	147.90	0.0000***
At most 1	150.60	0.0000***	65.06	0.0001***
At most 2	66.45	0.0001***	30.27	0.3504
At most 3	29.03	0.4108	25.81	0.5833
At most 4	29.10	0.4077	26.41	0.5503

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

6. Panel Dynamic OLS Regression Results for Prague

log(apartment	PDOLS	FMOLS	CCR	OLS
log(building plot price)	0.3852 (0.1580) **	0.2832 (0.1424) *	0.2632 (0.1412) *	0.3852 (0.0927) ***
unemployment	-0.0513 (0.0234) **	-0.0486 (0.0211) **	-0.0354 (0.0245)	-0.0513 (0.0137) ***
migration	0.0116 (0.0081)	0.0203 (0.0073) **	0.0275 (0.0097) ***	0.0116 (0.0048) **
age structure	-0.0280 (0.0102) ***	-0.0383 (0.0092) ***	-0.0437 (0.0103) ***	-0.0280 (0.0060) ***
REER	0.0098 (0.0037) ***	0.0122 (0.0034) ***	0.0128 (0.0034) ***	0.0098 (0.0022) ***
ECT	-0.16 (0.05) ***	-0.17 (0.05) ***	-0.18 (0.05) ***	-0.16 (0.05) ***
R^2	0.95	0.95	0.94	0.95
adjusted R^2	0.95	0.94	0.93	0.95

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

7. Cross-Sectional Dependence and Slope Homogeneity Tests

Table 13 presents the results of Breusch-Pagan LM test and Pesaran CD test for cross-sectional dependence. Based on the results, we can conclude that both tests rejected the null hypothesis of no cross-sectional dependence.

Table A8 Cross-Section Dependence Test

Test	Statistic	Prob.
Breusch-Pagan LM	6889.27	0.0000***
Pesaran CD	82.98	0.0000***

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

Table 14 shows the results of Pesaran-Yamagata test (or Delta test) for slope homogeneity. Under the null hypothesis, slope coefficients are homogenous. In such case, static panel estimators, such as fixed effects or random effects, are biased. In our case, the delta test rejected the null hypothesis, which suggests that the slope coefficient varies across time and that the analysis of our data requires a dynamic panel estimator.

Table A9 Slope Homogeneity Test

Test	*Statistic*	*Prob.*
Delta	*5.3550*	*0.0000****
Delta adjusted	5.7200	*0.0000****

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

REFERENCES

Asal M (2018): Long-Run Drivers and Short-Term Dynamics of Swedish Real House Prices. *International Journal of Housing Markets and Analysis*, 11(1):45-72.

Baltagi BH (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 (2018): Fundamental Determinants of Real Estate Prices: A Panel Study of German Regions. *International Advances in Economic Research*, 24:25-45.

Breitung J (2000): The Local Power of Some Unit Root Tests for Panel Data. *Advances in Econometrics*, 15:161-177.

Capozza DR, Hendershott PH, Mack C, Mayer CJ (2002): Determinants of Real House Price Dynamics, *NBER Working Paper*, 9262, October.

Cempírek P (2014): Housing Prices in the Czech Republic and Slovakia: Regional Comparison, Bachelor thesis, Charles University, Czech Republic.

Choi I (2001): Unit Root Tests for Panel Data. *Journal of International Money and Finance*, 20:249-272.

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

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

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

CZSO (2020): Ceny sledovaných druhů nemovitostí (Prices of Observed Types of Real Estate).

Dickey DA, Fuller WA (1979): Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 47:427-431.

Dröes MI, 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,* 49(3):367–388.

Engle RF, Granger CWJ (1987): Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55:251-276.

Granger CWJ (1969): Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*, 37:424-438.

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

Hlaváček M, Komárek L (2011): Regional Analysis of Housing Price Bubbles and Their Determinants in the Czech Republic. *Czech Journal of Economics and Finance*, 61(1):67-91.

Huynh-Olesen DT, 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. *OeNB Focus on European Economic Integration*, 2:52-76.

Im KS, Pesaran MH, Shin Y (2003): Testing for Unit Roots in Heterogeneous Panels. *Journal of Econometrics*, 115(1):53-74.

Johansen S (1995): Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. *Oxford University Press*, Oxford, United Kingdom.

Kalabiška R, Hlaváček M (2020): Regional Determinants of Housing Prices in the Czech Republic. *IES Working Papers*, 9/2020, IES FSV, Charles University.

Kao C (1999): Spurious Regression and Residual-Based Tests for Cointegration in Panel Data. *Journal of Econometrics*, 90:1-44.

Kao C, Chiang MH (2000): On the Estimation and Inference of a Cointegrated Regression in Panel Data. *Nonstationary Panels, Panel Cointegration and Dynamic Panels*, 15:179-222.

Levin A, Lin CF, Chu C (2002): Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties. *Journal of Econometrics*, 108(1):1-24.

Liew VKS (2004): Which Lag Length Selection Criteria Should We Employ? Economics Bulletin, 3(33):1−9.

Pedroni P (1999): Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors. *Oxford Bulletin of Economics and Statistics*, 61:653-670.

Phillips P, Hansen B (1990): Statistical Inference in Instrumental Variables Regression with I(1) Processes. *Review of Economic Studies*, 57:99-125.

Phillips P, Perron P (1988): Testing for a Unit Root in Time Series Regression. *Biometrika*, 75:335-346.

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

Saikkonen P (1991): Asymptotically Efficient Estimation of Cointegration Regressions. *Econometric Theory*, 7:1-21.

Schnure, C. (2005): "Boom-Bust Cycles in Housing: The Changing Role of Financial Structure". *IMF Working Paper* No. 05/200

Stock JH, Watson MW (1993): A Simple Estimator of Cointegrating Vectors in Higher Order Integrated System. *Economometrica*, 61:783-820.

Tsatsaronis K, Zhu H (2004): What Drives Housing Price Dynamics: Cross-Country Evidence. *BIS Quarterly Review*, March.

Xu L, Tang B (2014): On the Determinants of UK House Prices. *International Journal of Economics and Research*, 5(2):57-64.

Zemčík P (2011): Is There a Real Estate Bubble in the Czech Republic? *Czech Journal of Economics and Finance*, 61(1):49-66.

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