

**CHARLES UNIVERSITY IN PRAGUE**

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

**Peter Smolár**

**Regional Determinants of Housing  
Affordability in the Czech Republic**

*Bachelor thesis*

Prague 2016

**Author:** Peter Smolár

**Supervisor:** Mgr. Hana Hejlová

**Academic Year:** 2015/2016

## Bibliographic note

SMOLÁR, Peter. (2016) “The Brilliant Theory of Everything”, Bachelor thesis, Charles University, Faculty of Social Sciences, *Institute of Economic Studies*, Supervisor Mgr. Hana Hejlová.

## Abstract

The main motivation behind this study lays in identification of regional determinants in the Czech Republic. Because of lower manifested heterogeneity of apartment in comparison with houses, housing affordability was measured by a ratio of average buying price of an apartment over average annual disposable income per capita. Regional housing affordability was judged *en bloc* and was not measured in an attempt to identify subpopulations threatened by high housing costs. The object of interest was the impact of individual dependent variables on the price-to-income ratio. For the purpose of this study, only factors affecting at the same time housing prices and housing income, either directly or indirectly through correlation, are taken into consideration. The estimation of our model suggests that economically and statistically significant are the variables for the relative number of divorces, births or the proportion of urban population. All of the effects, disregarding only one variable, went in the same direction as they were expected to go. Also, it was possible to provide an answer to the question whether inclusion of composite determinants, such as natural population growth, may affect the interpretation of the results. In accordance with empirical studies on regional housing prices determinants, the study concludes that inclusion or exclusion of Hl. m. Praha region does not noticeably affect the results of the estimation.

## Abstrakt

Hlavná výskumná otázka spočívala v identifikácii regionálnych determinantov dostupnosti bývania v Českej Republike. Okrem nej sa spoločne spolu s ozrejmovaním skúmanej tematiky vynorili ďalšie, otázky doplnujúce. Kvôli menšej vykazovanej heterogenite bytov v porovnaní s domami, bola dostupnosť bývania meraná podielom priemernej kúpnej ceny bytu v danom kraji a priemerného čistého ročného príjmu na osobu. Dostupnosť bývania nebola posudzovaná pre jednotlivé typy domácností. Objektom záujmu štúdie bol vplyv jednotlivých nezávislých premenných na zmieny indikátor dostupnosti bývania. Za determinanty dostupného bývania sa pre účel tejto práce považujú veličiny ovplyvňujúce ako cenu bývania tak príjem domácností a to buď priamo alebo nepriamo prostredníctvom korelácie s iným determinantom. Premenné ktoré sa ukázali byť štatisticky a ekonomicky významnými boli napríklad relatívny počet rozvodov, narodení alebo relatívna početnosť ľudí žijúcich v mestách. Až na jednu výnimku, u všetkých nezávislých premenných prevládol ich očakávaný účinok na dostupnosť bývania. Podarilo sa taktiež zodpovedať otázku o možnom ovplyvnení interpretácie premenných za použitia zložených makroekonomických indikátorov, ako napríklad prirodzený prírastok populácie. Zhodne so závermi empirických štúdií o determinantoch cien nehnuteľnosti sa podarilo preukázať, že zahrnutie pozorovaní vyššieho samosprávneho celku Hl. m. Praha má na výsledky takmer nepozorovateľný vplyv.

## Klíčová slova

cena k príjmom, determinanty dostupnosti bývania, regionálny výskum

## **Keywords**

price-to-income, housing affordability determinants, regional analysis

**Range of thesis:** 65449 characters on 25 pages

## **Declaration of Authorship**

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature. The author hereby declares that all the sources and literature used have been properly cited.

The author hereby declares that the thesis has not been used to obtain a different or the same degree.

Prague, July 25, 2016

---

Peter Smolár

## List of contents

<b>1.</b>	<b>INTRODUCTION .....</b>	<b>7</b>
<b>2.</b>	<b>HOUSING AFFORDABILITY .....</b>	<b>8</b>
<b>3.</b>	<b>DETERMINANTS .....</b>	<b>9</b>
	<i>a. Housing prices .....</i>	<i>9</i>
	<i>b. Household income .....</i>	<i>10</i>
	<i>c. Housing affordability .....</i>	<i>11</i>
<b>4.</b>	<b>DETERMINANTS .....</b>	<b>19</b>
	<i>a. Data and variables .....</i>	<i>19</i>
	<i>b. Preliminary data inspection .....</i>	<i>21</i>
	<i>c. Model introduction .....</i>	<i>24</i>
	<i>d. Results of estimation .....</i>	<i>26</i>
<b>5.</b>	<b>CONCLUSIONS .....</b>	<b>30</b>
<b>6.</b>	<b>BIBLIOGRAPHY .....</b>	<b>32</b>
<b>6.</b>	<b>APPENDIX .....</b>	<b>35</b>

## **I. Introduction:**

After economic transition occurred in Czech Republic, a notable increase in costs of housing due to adaptation to market economy meant that the proportion of household expenditures on housing doubled between the years 1990 and 2003 (Sunega & Mikezsová & Lux 2009). What's more, this increase on global level was accompanied on the regional level by presence of significant housing affordability gaps. Later specified measure of housing affordability, price-to-income, showed that in the year 2014, affordability indicator in the region with least affordable housing was approximately equal to four times its value in the region from the opposite side of affordability spectrum. The scale of disproportion should make anyone attentive to readily available explanations. Due to country-specific heterogeneity of real estates, reflected in housing prices, conclusions drawn about housing affordability determinants outside Czech regions cannot be simply adopted without additional empirical evidence. Although some studies on the problematic of housing prices and housing affordability determinants on the regional level in Czech Republic exist, it is safe to say that housing prices, not to mention housing affordability, do not constitute what one would call traditional field of Czech economic research. To author's knowledge, due to lack of data, the research on owner-occupied apartment affordability has been so far done based on real-life statistics from which fictional standardized household types were created<sup>1</sup>. The estimations in this thesis are based on real-life panel data on regional scale. A particular source of inspiration was found in a study by Michal Hlaváček and Luboš Komárek (Hlaváček & Komárek 2009), which employed a similar econometric approach in regional housing prices determinants estimation. Rather than simply trying to reaffirm the obtained results, this paper tries to uncover determinants which are specific to housing affordability and are needed to be estimated anew. For this reason, the core of this thesis is formed by exhaustive discussion on theoretical and empirical literature which should justify the use of employed determinants and serve to forecast their expected effect on household income, housing prices and subsequently housing affordability.

The primary objective of this thesis is to fill the academic void by estimating the effects of potential housing affordability determinants on regional longitudinal panel data. In order to do so, a series of OLS estimations are run on the data from Czech NUTS 3 regions collected for 14 consecutive years. Besides the main task a few interesting and testable questions have arisen. First off, is there a statistical and/or economic difference between using a composite variable - such as natural population growth - or should such variable, when possible, be partitioned into its integral parts – in this example, rate of births and deaths. Other questions are based on the specific position of Hl. m. Praha in comparison with other regions. Namely secondly, does inclusion of the Hl. m. Praha region interfere

---

<sup>1</sup> See (Sunega & Mikezsová 2015) or (Sunega & Mikezsová & Lux 2009).

with the results obtained by first-differenced estimators and are these results still valid or should the inclusion of relevant dataset for the capital region be avoided in empirical studies on housing affordability. Lastly, are the determinants that are used to account for housing affordability satisfactory when it comes to explaining the variance in price-to-income ratio for Prague?

After this introductory, the study is organised as follows. Contents of the following chapter are formed by an overview of housing affordability measures. Once a measure for this study is selected, household income and housing prices determinants are debated in the third section. What must then follow is the empirical analysis. The study is complete with the discussion over results and finally a series of concluding remarks.

## **II. Housing affordability**

A common consensus in household affordability literature<sup>2</sup> classifies approaches to affordability analysis into three categories.

First off, the indicator measure of affordability is usually thought of as costs of housing to household income ratio. These vague terms are substituted by real-life statistics depending on the selection of subpopulation whose housing affordability is to be analysed. Once the subpopulation of interest is specified, threshold for “appropriate housing” is then established as a rule for judging values of the indicator. Common indicators of this type are so called rent-to-income or price-to-income measures of affordability in which either rent or price of real estate, in an appropriate unit or in total per estate, are related to household income.

Secondly, a so called reference method avoids using a threshold level and rather defines the relevant housing costs with reference to either other housing sectors or a subgroup of subpopulation which should be able to afford it. Lastly, residual income method looks at the income which households are left with after the housing costs are paid. Similarly to indicator measure, after subtracting the costs a threshold needs to be set for normative examination of housing affordability.

One of the main problems encountered in Czech research on housing affordability is the lack of data. To deal with this problem researches in the past selected higher NUTS classification and dealt with the subsequent loss of potential explanatory variables (Kostecký & Vobecká 2009) or obtained their data by deriving fictional standardized households from relevant real-life statistics<sup>3</sup>. Even though in this thesis a severe lack in officially published data on mortgage repayments forced an abandonment of the intended residual approach to housing affordability, none of these approaches was adopted because the objective of this paper differs.

---

<sup>2</sup> See (Kostecký & Vobecká 2009), (Sunega & Mikezsová 2015) or (Sunega & Mikezsová & Lux 2009).

<sup>3</sup> This approach was used in (Sunega & Mikezsová 2015) and the Framework developed in (Sunega & Mikezsová & Lux 2009)



As the name of this study suggests, debate over housing affordability is not to be held on the affordability of housing for various types of households *per se* but rather consists in regional analysis of affordability determinants for all the households in a set region. No thresholds in the normative measures of affordability are needed to be adopted since only the effects of selected variables on housing affordability rather than evaluation of housing affordability for selected types households is of interest. Some of the criteria upon which the households were standardized in abovementioned studies— such as age, education or employment status of members of each household type - are incorporated in their aggregate form to provide a better *en bloc* analysis.

The selected measure of housing affordability is price-to-income ratio. Even though it is widely used to measure affordability of owner-occupied apartments (OECD 2005), it is certainly not devoid of shortcomings since it does not reflect interest rates on mortgages nor obtained loans for that matter (Kostelecký & Vobecká 2009). The measure is also expected to be upwardly biased, since the average income of households who decide to buy rather than to rent an apartment will very probably be higher than the global average income. (Hrachovec 2012) A technical definition of housing affordability is developed in the chapter dedicated to data and variables.

### **III. Determinants:**

This chapter sets for aim to clarify why proposed determinants are to be used, explore potential links of correlation between them and provide a discussion about their expected effects on household income, housing prices and affordability. The proposed determinants are either of demographic or economic character and are classified accordingly. A common framework in Czech research on housing prices determinants exists and is therefore reviewed before discussing the heterogeneous mass of research on household income determinants. As is explained, only factors affecting both housing prices and household income are of interest for this study.

#### **a. Housing prices determinants:**

An increase in general fear of a potential housing market bubble in Czech Republic, caused by the latest major financial crisis of 2007-2008, fostered development in Czech housing prices literature which was to that date basically non-existent. P/I ratio subsequently became quite popular as a tool for rough estimations of housing market discrepancies. In the relevant Czech literature on the subject of housing prices, only (Égert & Mihaljek 2007) do not mention named indicator.

First found to be mentioned in (Égert & Mihaljek 2007), who in turn used a framework provided in (HM treasury 2003), a common equilibrium model of housing prices is used by the majority of studies by Czech authors. The mentioned theoretical approach, also adopted by this study, sees equilibrium on the housing market as equilibrium between housing demand and housing supply. Each of these sides is

in turn determined by a specific set of factors which helps with preliminary selection of determinants and consequently justifies their use.

Supply in this model depends on housing prices and then on the real costs of construction which can be taken as a function of residential plot prices, construction workers' wages and the cost of material. (Égert & Mihaljek 2007) Construction costs are reflected into housing prices only with a long lag. (Hlaváček & Komárek 2009) Also as discussed earlier, neither the residential plot prices nor cost of material are necessarily connected with housing prices as well as disposable income. For these reasons only the wage in construction is considered as a household affordability determinant.

Demand for housing is also naturally dependent on housing prices, as well as other factors such as household income, the interest rate on housing loans, financial wealth of households or demographic and labour market factors (Égert & Mihaljek 2007).

A positive relation between household disposable income and housing prices is established on the theoretical grounds of the equilibrium model and on empirical estimations<sup>4</sup> done on real-life data in the mentioned literature. Determinants of price-to-income ratio therefore have to be estimated separately if they are believed to affect both the housing prices and household wealth. Without doing so, one could never be sure of the overall impact of the proposed determinant on price-to-income ratio.

For the remaining factors, which affect either solely the variable in numerator or denominator of P/I ratio and are not thought to be interdependent with the included variables, the analysis is redundant as nothing should change about their expected effect on either housing price or income, once they would be included into the model.

Most of the demand for housing can be explained by disposable income whilst many of the remaining demand factors are affecting disposable income directly or indirectly (Hlaváček & Komárek 2009). This stresses the need to estimate housing affordability and housing prices separately, especially when the same study (Hlaváček & Komárek 2009) asserts that housing prices in Czech regions are primarily demand driven.

#### **b. Household income determinants:**

Literature on household income is mostly centred around inequality at the level of individual households. Making use of determinants commonly perceived by literature as significant in explaining the variance in household income therefore poses a challenge.

---

<sup>4</sup> proxied by wages or GDP growth Dipl GDP growth,, Hlaváček wages, GDP growth 1

Aggregation of these microeconomic determinants seems viable when it is reasonable to assume that the impact of individual differences in determinants on household income when aggregated will manifest itself similarly on the regional scale. Such is the case for education, employment and unemployment or marital status. Mentioned factors are commonly recognised as determinants of inter-household inequality and can easily be transformed into their aggregated form.

Other determinants simply cannot be aggregated without justification. Such is the case of education which is consistently found to have a positive impact on real wage of individuals. However when aggregating on the regional level the possibility of over-education needs to be accounted for. There also remains a subgroup of household income determinants unsuitable for aggregation on the regional level since doing so would not make any sense as differences on the level of households cease to exist once the data are aggregated. Of this last subgroup, gender is a perfect example.

As wage, in itself largely regionally varied, is one of few integral parts of household income, understanding the possible causes of variation in wages should lead us at least partly to an understanding of regional household income differences. Therefore to an extent, articles on regional wage differences will be used to argue for inclusion of some of the determinants. For the purpose of this work, an assumption that determinants of wage affect household income similarly will be held.

### c. Housing affordability determinants:

#### **Age (demographic determinant):**

In this study, participation rate and unemployment rate are divided into two categories based on age. The variables for various age groups are used as a proxy to experience of workers. Experience as a part of human capital appears fairly often in research papers on wage inequality<sup>5</sup>. Age is a commonly used proxy for experience<sup>6</sup>. In fact, age reflects so many different factors affecting wage that if one were to distil the pure effect of age on earnings, a use of longitudinal micro-economic data on worker's experience, education, gender and others would be required (Luong & Hébert 2009).

As to the relation between age and wages, current situation in Czech Republic reflects a few differences between public and private sectors. In the private sector, a substantial average gross wage gap of 28.7% exists between the workers in 20-29 and 30-39 age ranges. After reaching the latter age range, where worker's wage is equivalent to 29642 CZK per month and roughly 106.7% of total the average, the average gross wage continues to plateau and eventually falls by about 3% for workers above 60. In the public sector, a similar gap of about 21% exists between people in their 20ties and 30ties. However, the average monthly wage continues to rise with seniority and is the highest for people above 60 when it is at the level of 30588 CZK per month, which represents a 9.4% bonus to the

<sup>5</sup> For example, see (López-Bazo & Motellón 2012),(Cahill & Gager 2014) or (Ciešlik *et al.* 2013)

<sup>6</sup> See statements by (Luong & Hébert 2009) or ( Earle 2010)

overall average<sup>7</sup>. Accordingly, it is expected that higher portion of middle-aged workers in the population translates to higher average real wages.

With regards to the rising popularity of tertiary education amongst people born after 1980 and with relatively lower educational levels of elderly in comparison with states of Northern and Western Europe (Bakalová & Boháček & Münich 2015), patterns connecting different age subpopulations with varying levels of education are expected to emerge.

As for the younger economically dependent subpopulation, economic consequences of parenthood are analysed in the subchapter dedicated to marital status and parenthood. Also the relation between age and unemployment is explained in the subchapter dedicated to unemployment.

The effects of the age composition of a population on housing prices are expected to be varied depending on the respective age groups (Hlaváček & Komárek 2009). A population boom is expected to strengthen the incentives for independent housing. Respectively, a higher number of people in the age typical for household formation should manifest itself in an increased demand for housing (OECD 2005).

On the other hand, the number of deceased acts in the opposite direction - fuelling the housing supply and hence lowering the equilibrium price. However due to the legal aspect of hereditary succession and potential reconstructions it is important to account for a certain time lapse before the vacated apartments re-enter the market.

#### **Average gross wage in construction (economic determinant):**

For wage in construction, there is not much explanation needed to understand why it was included. The effects on household income are straightforward consequences of applied premise about relation between the former and wages. On the other hand it is the only housing supply factor included. It is expected to affect housing prices with a positive sign but to better see this lagged variables should be specified. No apparent links with other included variables seem to exist.

#### **GDP per capita dynamics (economic determinant):**

The inclusion of growth of GDP per capita is based on a theoretical body of economics which suggests a cyclical character of real wages in relation to GDP. When it comes to cyclicity of real wages, a fact that emerges quickly is that economic theory is not uniform on the subject. Not only a division between equilibrium and Keynesian models but also an internal divide in these respective groups arise on the subject of pro-cyclicity or counter-cyclicity of wages. (Abraham & Haltiwanger, 1995) What may be worse is that empirical studies resulted in more of a confusion than a conclusion.

---

<sup>7</sup> Computations done on data provided by MPSV

The aforementioned article by (Abraham & Haltiwanger, 1995) tries to uncover why different studies concerning the US at the national level offer different results. Their conclusions lead us to doubt the systematic stability of cyclical character of real wage in the US. This denial is accompanied by a collection of factors determining the results of empirical studies such as the time period used for the study, the choice of a price deflator or the composition of industry.

Leaving the borders of USA and concentrating on manufacturing sectors of 18 OECD countries, (Messina, Strozzi & Turunen 2009) first confirms that the methods and data applied affect the results of a study. Once these factors are controlled for, it is established that some of the selected countries follow the pro-cyclical or counter-cyclical pattern whilst others follow a-cyclical pattern, in relation to real wages. (Abraham & Haltiwanger, 1995)

GDP per capita can also serve as a proxy for wealth, which is considered to be one of housing demand factors (1). The effects of growth of GDP per capita on housing prices are therefore expected to be positive. This means that whether growth of GDP per capita is a decisive factor depends on the cyclical nature of household income.

#### **Education (demographic determinant):**

On micro-economic real life data there is a well-established positive correlation (Card 1999) between higher levels of education and higher real wages. Although it may seem tempting to go a step further and make causality out of correlation (Card 1999) some scholars oppose this.

The strong, clear cut positive effect of education on real wages does not transfer to the macro-scale in its full extent. (Psacharopoulos, 2004) The macro-economic effect is still positive but its absolute size is largely varied. What's more is that this variation remains inexplicable (Trostel, 2002). Moving to regional scale through academic void brings along risks - namely the risk of not accounting for over-education. A recent study concerning Ukraine points to its socialist heritage as a plausible cause of over-education (Kupets, 2016).

Using a normative measure, it has been estimated that the total 39.7% of employees aged 15–70 are over-educated. The most overeducated for their working position are absolvents of so called učiliště, described by the author as "Soviet-type" college. However a recent study on Czech data (Laslopova 2015), using the same normative approach, estimates the total over-education of about 6.2%. Such low total results are also obtained in (Mysíková, 2014)

Apart from expected positive direct effect of higher education on average real wages, the education variable will be included because of apparent links with other variables, such as unemployment, migration or regional GDP. Except for GDP, these links will be discussed in later subchapters.

Traditionally, education can be linked to GDP through Solow Model of growth with human capital (Mankiw, 1992). Effects on GDP, although positive, vary largely depending on country and these differences are not easily explainable (Pritchett, 2001)

The link between education and housing prices is not a direct one. Education, like unemployment, is thought to affect housing prices through household income. (Hlaváček & Komárek 2009) The reason why it is then included is because as is shown, it is more than likely correlated with other variables which affect both housing prices and household income.

### **Marital status and parenthood:**

Marital status appears fairly often as a dummy variable throughout micro-economic studies on wage inequality (for example see Cahill & Gager 2014), (Lopez Bazo & Montellon 2012) or (Cieslik & Rokicki 2013). Due to specifics of legislature and social norms, the focus shall lay on the findings about Czech Republic.

Between the years 1993 to 2014, Czech Republic has seen a steady increase in portion of the children born outside of a marriage from 12.7% to 46.7%.<sup>8</sup> Yet it cannot be disputed that one of the main purposes of matrimony is the conception and bringing up of children. Economic consequences of such "matrimonial task" as parenthood are laying unevenly on sexes. These consequences have been recently analysed in a study by Mariola Pytlíková: Wage Differences Related to Motherhood and Children in the Family. (Pytlíková 2015)

It shows that, only around 1% of men make use of the long (in comparison with other European countries) parental leave guaranteed by law. The difference between the average number of hours worked weekly by men and women, which is 2 hours for childless men and women, increases to 6.4 hours after birth of a child. This difference returns to a value of 2.5 hours once the child reaches adulthood. Participation of women between 25-49 drops from 90% to 20% after a child is born. The overall effect of parenthood on parent's wages and hence household income may therefore be positive or negative, depending on whether the hours worked by the father compensate for the loss undergone by the mother.

As to the relation between unemployment and parenthood, another study from the series (Bičáková & Kalíšková 2015) elaborates on the subject. Partly because of absence of preschools for children below 3 years of age, partly because of social norms which emphasises the bond with young children, women are taking the three years long parental leave. As the time passes, productivity of these women falls. This is due to technological changes which happen in their absence, stagnation of human capital and further absence caused by temporary sickness of the child.

---

<sup>8</sup>Computations done from the data by CSO

The risks of unemployment for the subpopulation of men and women whose child is older than 15 are very close (namely 4.9 % for women and 4.2 % for men). An introduction of a child below 15 changes these risks. With women, it increases substantially to 6.4 %, whereas men see a reduction to 2.8 % which can be explained by an increase in their motivation to work and avoid unemployment. Education, part of human capital, reduces these risks for women.

As to divorce, a study by Jiří Šatava and Petr Janský (Šatava & Janský 2015) points to similar roots of its negative economic consequences. Firstly, when it comes to gross disposable income, households whose head is divorced and is living without a partner make for 12.6% of all households. However, their gross household income makes up only 8.5% of the total. Secondly, Division of labour amongst the married, even if speculatively beneficial in itself, makes one of the pair specify in an occupation outside the labour market. Such division halts the accumulation of human capital for one of the partners and brings along the already mentioned higher risk of being unemployed.

Divorces, newly formed marriages and the already mentioned population growth are thought of as being demographic determinants of housing prices pushing on demand for housing in the same positive direction. From this set of factors, divorces would seem to have the biggest influence on housing prices as the need for new housing is the most imminent and understandable for divorced couples. Accordingly, the only study on Czech regions covering all of these factors concludes (hlaváček) that whereas population growth and divorce rate are found to be statistically significant the same is not true for marriage rate. To compare the economic significance of the number of births and divorces, the variable for population growth has to be partitioned.

### **Migration (demographic determinant):**

Early economic theory on migration, as seen by Hicks, argues for differences in real wages among regions as the main migration motivator. (Dickey 2015) Workers from different regions are seen as substitutes. A region with relatively higher real wages would see an influx of workers which would *ceteris paribus* consequently lower the average real wages in the region of arrival and higher the average real wages in the region of departure. (Pouliakas, K., D. Roberts, E. Balamou, & D. Psaltopoulos 2009) (Dickey 2015)

This equalising effect on average wages has been tested for and found. However, apart from the studies concerning the US, the economical if not statistical significance of these findings is negligible.<sup>9</sup> For the Canadian case, based on (Wrage 1981), it would take 130 years to close a wage gap of 10 percent. The expected positive effect of emigration on average wage is not found.<sup>10</sup>

---

<sup>9</sup> see literature overviews in (Fendel 2016) or (Pouliakas, K., D. Roberts, E. Balamou, & D. Psaltopoulos 2009))

<sup>10</sup> The author concludes by: "As an earning adjustment mechanism, migration simply does not work very well in Canada."

Another study from Germany, points to an overall positive effect of immigration on average wages. When the foreign influx of migrants has been controlled for and only the effect of inter-regional immigration was tested for, the study reported a negative effect of immigration on average wages. (Fendel 2016) Although these results may seem counterintuitive at first, they can be explained by and hence strengthen the case for the human capital model of migration.

By incorporating human capital and hence relaxing the assumption of homogeneity of the migrating workforce, we can now explain why the overall effect of immigration was positive when the foreign workers were accounted for. A new theoretical assumption, strengthened by empirical research, arises that migrants with higher levels of human capital are *ceteris paribus* more prone to migrate. (Dickey 2015) In an oversupply of workers low on human capital, the "skilled-wage premium" is a high motivator to foreigners with high human capital. (Dickey 2015)

Whereas an influx of workers with comparable levels of human capital should lead to a decrease of average wages, the same is not always true for those who enter with higher levels of human capital. Higher education, consistently with this new model, is found to increase mobility. Evidence for this statement lays in two studies (Wiess 2015) and (Machin, S., K. G. Salvanes, & P. Pelkonen 2012) done on countries which introduced additional obligatory years of schooling.<sup>11</sup>

Internal migration between 2001 and 2004 in Czech Republic follows this education-migration pattern where, relative to their representation in population, people with at least complete secondary education migrate more.<sup>12</sup> In 2001, people with tertiary education from 20-34 age range accounted for most migrations, relative to the rest of migrating population with different levels of attained education. With regards to the aforementioned and to the rising popularity of tertiary education (as already mentioned in the subchapter dedicated to age) we expected the overall effects of immigration on average real wages to be positive.

When it comes to unemployment, a persistent fear adjacent to early theory on migration is that immigration in combination with rigid wages structure will result in higher overall unemployment in the region of arrival. This effect of migration on non-migrant employment/unemployment has been proven to be minimal or non-existent (Dickey 2014) by empirical analysis.

As to the propensity of a worker to migrate, study of Šatava & Jánký provides a useful insight.<sup>13</sup> First of all, the probability of migrating is higher for an unemployed worker. This idea was also articulated

---

<sup>11</sup> (Machin, S., K. G. Salvanes, & P. Pelkonen 2012) has been conducted on the data from Norway, (Wiess 2015) has been conducted on 8 countries and confirmed the results of (Machin, S., K. G. Salvanes, & P. Pelkonen 2012)

<sup>12</sup> From this period, it no longer possible to follow internal migration with regards to education due to methodological shift (see [https://www.czso.cz/csu/czso/13-1135-08-2001\\_az\\_2007-1\\_2\\_ucastnici\\_vnitriho\\_stehovani\\_v\\_cr](https://www.czso.cz/csu/czso/13-1135-08-2001_az_2007-1_2_ucastnici_vnitriho_stehovani_v_cr))

<sup>13</sup>(Šatava & Jánký 2015)



by (Wrage 1981) when it didn't observe the expected increase in average wages in the region of departure which should have occurred according to the early theory on immigration. Second of all, a worker is more prone to migrate if the region of departure has a relatively higher unemployment levels. Lastly, "migration propensities" are lower when the national overall unemployment rates are higher.

In the earlier cited study (Pouliakas, K., D. Roberts, E. Balamou, & D. Psaltopoulos 2009), immigration was found to have only a very small impact on GDP per capita, when regarded through the scope of earlier theories. The findings of another study dealing with the relation between GDP and migration are surprisingly similar to those obtained in (Fendel 2016). In the case of Sweden, convergence in GDP between regions occurs due to internal migration. The opposite is true for Norway. The author of the paper argues that convergence was achieved because of higher human capital variations between counties, among which the migration occurred.

As a demographic driver of housing demand, net migration is expected to be affecting housing prices positively. A theoretically established argument is that an increase in net migration, along with other factors such as natural increase in population, inflates the number of housing market-participants (HM Treasury 2003). Such pushing on the demand side of the housing price equation shall *ceteris paribus* lead to a higher equilibrium price - a conclusion strengthened by empirical evidence from Czech regions (Hlaváček & Komárek 2009). The effects of migration on housing prices are thought to be stronger than those on household income since not all migrants must be workers but all migrants need housing.

#### **Unemployment (economic determinant):**

For those unemployed, the part of household income formed by their wage is replaced by presumably lower state-aid. For those who live in a region affected by unemployment, an empirical pattern known as "Wage-curve" explains the lowering of their wages.

Wage-curve was established based on a research paper (Blanchflower & Oswald 1990). Throughout the years various rationales for the existences of wage-curve have been proposed (Blanchflower & Oswald 1995). The pattern, although still controversial in the US, has been recognised in at least 40 other countries, including Czech Republic. (Blanchflower & Oswald 2005)

The study by (Galusčák & Munich 2003), which is based on data from 1996-2001, suggests that wage-curve can be observed in Czech regions, but not without conditions. Factors that come into play when pondering the existence or absence of wage-curve are overall levels of unemployment, temporal aspect of unemployment and the share of public sector in the region.

There is not enough statistical evidence for wage-curve in regions with high employment in public sector, since in the public sector decision about wages is not done on the basis of unemployment. Also, the wage-curve is weak or absent in regions with high overall or long-term unemployment. The author proposes that in these cases the generous welfare system may provide a floor for the wage adjustment mechanism.

Up until 2008, the risk of unemployment concerned mostly the elder part of population. The financial crisis meant that not enough new working opportunities were created which induced high unemployment rates for younger labour market participants (CSO 2012).

Quoting the same analysis by ČSÚ: “The highest attained education plays an important role in determining the possibilities of employment for the younger population. It shows itself clearly that the unemployment rate is substantially dependent on the level of attained education - the higher level of attained means lower unemployment rate.”<sup>14</sup>

Taking a peek on data published by CSU, throughout the years 1993-2014, we see a similar education-unemployment pattern throughout the population. The average rate of unemployment for the subpopulation with no or only elementary education for the period is 21.25%, with the maximum of 28.48% reached in 2012. The relative numbers for the subpopulation with a university diploma are 2.37% for the average and 3.35% for the maximum.<sup>15</sup>

Unemployment is one of the factors affecting housing prices indirectly, channelling its effect through household income (Hlaváček & Komárek 2009). Loss of steady income can force households to reevaluate the decision of obtaining an apartment of their own. Although a plausible link of substitution between renting and buying an apartment exists on national level (Hlaváček & Komárek 2009), when considering convergence of apartment rents and prices quite a few time periods should be accounted for. For this reason unemployment is viewed as affecting the housing prices only indirectly.

### **Urbanization (demographic factor)**

For urbanized areas, relatively higher productivity of workers and consequently higher real wages are expected due to aggregation knowledge spill overs or better matching of skills (Cahill and Gager 2014). That is why the expected impact of urbanization variables on household income is thought to be positive.

---

<sup>14</sup> From original: “Nejvyšší dosažené vzdělání hraje významnou roli v možnosti uplatnění mladých osob na trhu práce. Zcela jednoznačně se ukazuje, že míra nezaměstnanosti je významně závislá na stupni dosaženého vzdělání, čím vyšší je míra dosaženého vzdělání, tím nižší je míra nezaměstnanosti.”

<sup>15</sup> Calculations based on data provided by CSO

As to the correlation with other determinants, migrants are expected to select towns over villages since urban areas because of better wages, overall employment opportunities and other factors such as greater anonymity. (Pouliakas, K., D. Roberts, E. Balamou, & D. Psaltopoulos 2009)

Urbanization is also used as a proxy for the size of municipalities. This variable among other few is accounted for when housing price index is calculated and rightly so since for the year 2014, apartment prices in municipalities with total inhabitants below 2000 accounted on average only for 56.83% of those in other municipalities combined. The effect of higher proportion of urban population on housing prices is therefore expected to be positive and the overall effect on housing affordability depends on the

#### **IV Empirical analysis:**

##### **a Data and variables:**

All the data used in this study were obtained from publications available online at the official CSO website<sup>16</sup>. They were collected in a way to form a panel data structure. Cross-sectional dimension of this panel is defined by NUTS 3 Czech regions and the time dimension is defined by years 2001 to 2014. The data were collected on yearly basis. First differencing estimators were selected over fixed or random effects estimators because they appear to fulfil the assumptions needed for asymptotic consistency better.

The data on housing prices come from periodical publication named “Real Estate Prices”<sup>17</sup>. As the focus of his study is to analyse price-to-income ratio, average buying prices per and average size in appropriate units are the statistics of interest.

When discussing affordability of housing, this study uses apartment related data. Such approach is taken because apartment prices are expected to be subject to less heterogeneity in comparison with house prices. Even though CSO controls for heterogeneity in a region in a specific point in time before incorporating an estate into the sample<sup>18</sup>, the temporal aspect of heterogeneity is not taken into account when calculating the average buying prices. Houses as opposed to apartments manifest a lower level of homogeneity throughout time. Temporal heterogeneity can be shown on quantifiable data<sup>19</sup> and could also be partly accredited to unperceived qualitative differences between sold houses. In the absence of variables capable of capturing such differences which can be potentially correlated to other

<sup>16</sup> The official site was (23.7.2016) accessible at: <https://www.czso.cz/csu/czso/home>

<sup>17</sup> From the original: “Ceny sledovaných druhů nemovitostí“, available online (23.7.2016) for 2001 and further at: <https://www.czso.cz/csu/czso/ceny-sledovanych-druhu-nemovitosti-2001-2003-n44phfdnvw>

<sup>18</sup> See any introduction to “Real Estates Prices”, explaining applied methodology.

<sup>19</sup> For example, the difference in sizes of sold apartments between the years 2001 and 2014 is virtually non-existent on national level, whereas the similar measure for houses has seen a 10% increase. (authors calculations)

explanatory variables, one runs the risk of introducing a bias caused by omitted variable. It is therefore best to avoid using prices of houses as a measure of affordability.

An alternative to using apartment buying prices would be to use index of apartment prices, calculated by CSO, to approximate the apartment price development and then set a base year from which the approximated apartment buying prices would be calculated. Such a step would effectively increase the total number of observations, since the housing index is readily available from the year 1999 whereas the data on apartment prices are provided online from 2001 onwards. Moreover, due to the construction of the index, some of the temporal heterogeneity disregarded by our model would be controlled for. Nonetheless, as the price index does not copy the development of buying price with sufficient precision, the selection of the base year for apartment prices would be subject to a fair amount of arbitrariness and would very probably cause a distortion in results.

As for the income part of proposed price-to-income, the yearly data on disposable income per capita obtained from regional accounts<sup>20</sup> will do. Average disposable income per households would arguably be a better measure since the demand for housing reflects decisions of households rather than those of individuals. Due to a lack of data necessary for construction of such statistic for the complete timespan between 2001 and 2014<sup>21</sup>, which delimits the availability of relevant apartment statistics, additional observations are going to be preferred over solely potentially better income statistic.

For all the above mentioned reasons the measure of housing affordability is represented by *afford* variable and for a certain time and region computed as follows:

$$afford = \frac{\text{average buying price of } m^2 * \text{average size of an apartment in } m^2}{\text{net disposable income per capita}}$$

Following the affordability measure of the proposed model are demographic variables on natural increase in population, births, deaths, net migration, immigrants, emigrants, divorces and newly formed marriages. Relevant statistics were all measured per 1000 inhabitants and were obtained from datasets of long-term regional development in selected indicators<sup>22</sup>. Their variable abbreviations are, in order: *nat\_inc*, *births*, *deaths*, *net\_mig*, *immig*, *emig*, *divor* and *marr*. In consistence with the provided discussion on determinants, *deaths\_lag* presents the *deaths* variable with one year lag. Since 2010, CSO no longer collects data discriminating between internal and external migration of citizens and foreigners. Also CSO does not continue in differentiating migrants based on the criterion of the

---

<sup>20</sup> See (23.7.2016):

<http://apl.czso.cz/pll/rocenka/rocnkavyber.volba?titul=Ukazatele%20v%20region%20E11n%20EDm%20E8len%20ED&mypriznak=RC&typ=2&proc=rocenka.presmsocas&mylang=CZ&jak=4>

<sup>21</sup> The other potential measure of disposable income is being measured

<sup>22</sup> The data could be accessed (23.7.2016) at: [https://www.czso.cz/csu/czso/regionalni\\_casove\\_rady](https://www.czso.cz/csu/czso/regionalni_casove_rady)

highest attained level of education. Under these conditions, not many implications of human capital theory on migration can be tested for.

From the very same source came the statistics on GDP per capita, proportion of urban population and average gross wage in construction. After some basic calculations, the variable representing the increase in GDP per capita in reference to last year values, *inc\_gdp\_pc*, was derived. Average gross wage in construction in current prices is presented in *constr\_wage* and the relevant variable for the percentage of the population living in towns is *urban\_pop*.

The basis for statistics on participation, unemployment and different highest attained levels of education is captured by 1993-2014 labour market time series data<sup>23</sup>. Participation rate is calculated for the population in their 30<sup>th</sup> to 50<sup>th</sup> year of life in order to reflect the experience induced wage premium associated with this age group. The variable *parti\_30to50* is accompanied by *parti\_rest* in order to cover every individual older than 15. The unemployment statistics, *unem\_30to50* and *unem\_rest* follow the same logic. Highest attained education is accounted for in *prop\_prim*, *prop\_sec* and *prop\_ter* which consecutively represents proportion of population over 15 years of age whose highest attained education is primary or no education (0 to 2 in ISCED classification), secondary education (3 and 4 ISCED) and finally tertiary education (5 and 6 ISCED).

#### **b. Preliminary data inspection:**

Before moving to the model specification and estimation, first a brief summary of statistics in use is provided.

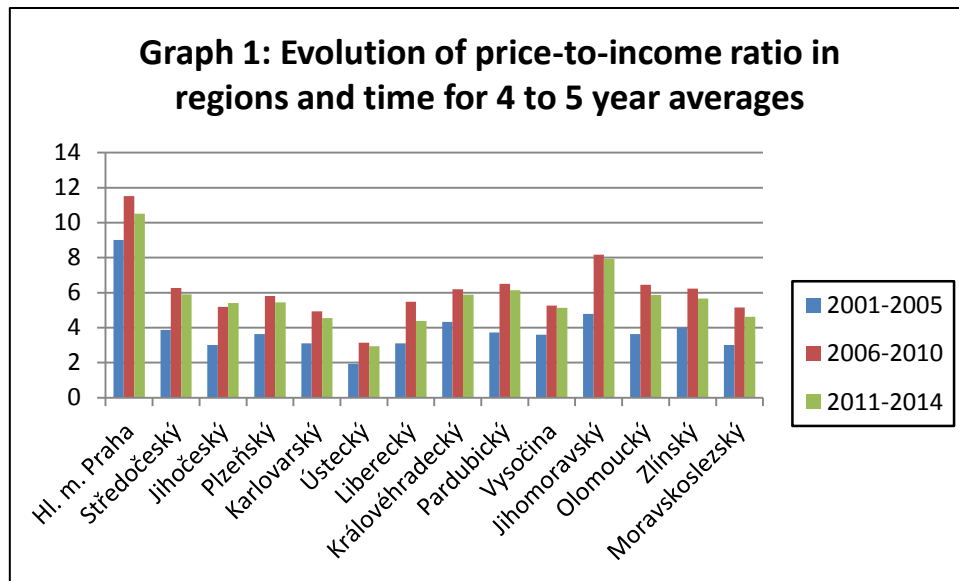
Shortly after economic transition, a rapid but regionally uneven growth of rental and housing prices occurred (Sunega & Mickszova & Lux 2009). The most notable gap arose between the capital and rest of regions. Both (Sunega & Mickszova 2015) and (Sunega & Mickszova & Lux 2009) assert that this period of growing inequality among regions lasted up to 2003 when it was replaced by a period of convergence lasting at least to 2006.

In order to ascertain further development of housing affordability, a commonly used measure of variation is constructed. In a ratio, standard deviation of price-to-income is compared to its time relevant mean. After that, it is possible to ascertain that the downwards trend continued – if the variation among regions is quantified as being equal to 44.37 % between the years 2000-2005, for the period between 2011 and 2014 it is equal to only about 31.82%.

Even if the convergence in housing affordability seems to exist among Czech regions, in present time the gap is still very much present. To better visualize, a simple graph relating P/I ratio to different time periods and regions follows.

---

<sup>23</sup> Data to be found (23.7.2016) at : <https://www.czso.cz/csu/czso/trh-prace-v-cr-casove-rady-1993-az-2014>



Another basic summary of employed statistics is presented by the following table. The same method applied to compute variability of P/I ratio is applied on the relevant statistics when computing variability in time and among regions. Whilst high temporal variability is desirable in order to account for the information loss induced by estimating first differences OLS, variability across regions is crucial for the explanatory power of the estimation

Statistics Variables	Variability		Average		
	in time	in regions	overall	highest values across time	lowest values across time
Afford	0.3207	0.2262	5.24	10.33 / Hl. m. Praha	2.65 / Ústecký
nat_inc	7.8237	3.1917	-0.12	0.57 / Liberecký	-0.78 / Zlínský
Births	0.0342	0.0751	10.18	10.91 / Středočeský	9.53 / Zlínský
Deaths	0.0239	0.0277	10.26	10.83 / Ústecký	9.72 / Jihomoravský
net_mig	1.133	1.679	1.88	10.9 / Středočeský	-1.55 / Moravskoslezský
Immig	0.23	0.5741	11.31	23.07 / Hl. m. Praha	4.19 / Moravskoslezský
Emig	0.0804	0.4428	9.41	12.22 / Hl. m. Praha	5.74 / Moravskoslezský
Marr	0.0883	0.0434	4.76	5.27 / Hl. m. Praha	4.46 / Zlínský
Divor	0.0795	0.1036	2.91	3.45 / Karlovarský	2.33 / Vysočina
urban_pop	0.0045	0.1754	69.23	82.06 / Karlovarský *	53.9 / Středočeský
Unem	0.1343	0.3276	6.98	11.39 / Ústecký	3.24 / Hl. m. Praha
unem_30to50	0.1139	0.3354	5.87	9.82 / Ústecký	2.74 / Hl. m. Praha
unem_rest	0.1416	0.3255	7.43	12.02 / Ústecký	3.46 / Hl. m. Praha
constr_wage	0.0865	0.1665	18420. 39	23717.79 / Hl. m. Praha	16762.02 / Karlovarský
prop_prim	0.0859	0.162	0.43	0.24 / Karlovarský	0.11 / Hl. m. Praha
prop_sec	0.1383	0.1557	0.19	0.67 / Hl. m. Praha	0.37 / Ústecký
prop_ter	0.2164	0.3895	0.11	0.26 / Hl. m. Praha	0.07 / Ústecký
gdp_pc_growt h	0.8115	0.1696	0.04	0.05 / Moravskoslezský	0.03 / Karlovarský

\* Disregarding Hl. m. Praha

When comparing regional statistics based on their total average, variables for Hl. m. Praha appears 12 out of 18 on the extremes as being either highest or lowest. This forces one to question the extent to which the proposed variables are capable of explaining the housing affordability for Prague. Also another question is whether for the remaining regions the estimations are sensitive to inclusion of the capital into the dataset. Answers to these questions are given in the post-estimation discussion over results.

**c. Model introduction:**

A panel of size 14 in the cross-sectional and 14 in the temporal dimensions, formed by the collected panel data, is analysed by pooled ordinary least square estimation using first-differenced variables. Estimation done in this manner rids the model from time-invariant omitted variables (Wooldridge 2013). Therefore to an extent, estimating first differences helps to prevent endogeneity of errors from occurring.

In order to save some space and render the model more intelligible, a “fd” prefix before a variable means that the variable had been first differenced, *i.e.*  $N$  being the size of the panel in cross-sectional dimension and  $T$  in temporal dimension, for a  $\{x_{it}\}_{i=1;\dots;N,t=1;\dots;T}$  stochastic process:

$$fd_{x_{it}} = x_{it} - x_{i(t-1)} \text{ for } i \in 1; \dots; N, t \in 1; \dots; T$$

Two models which serve as a basis for estimations are included in this paper. There is an expanded model, which partitions *nat\_mig* into *births* and *deaths*, *net\_mig* into *immig* and *emig*. Similarly, *parti\_rest* and *parti\_30to50* variables are broken into *unem\_rest* and *unem\_30to50*. Motivation behind having two estimations is to provide an insight firstly on the variable-sensitivity of obtained results and secondly on the dominant determinant, without being concerned about the absolute values of coefficients and rather focusing on their sign. After that, estimations in which Hl. m. Praha region is included in the sample are done separately from those done on sample without the capital in order to establish, whether the results are volatile in this aspect.

The restricted (1) and expanded (2) models are constructed in the following manner:<sup>24</sup>

(1):  $fd\_afford\_det_{it}$

$$\begin{aligned} &= \beta_0 + \beta_1 fd\_nat\_inc_{it} + \beta_2 fd\_net\_mig_{it} + \beta_3 fd\_marr_{it} + \beta_4 fd\_divor_{it} \\ &+ \beta_5 fd\_urban\_pop_{it} + \beta_6 fd\_unem_{it} + \beta_7 fd\_constr\_wage\_det_{it} \\ &+ \beta_8 fd\_prop\_prim_{it} + \beta_9 fd\_prop\_sec_{it} + \beta_{10} fd\_inc\_gdp\_pc_{it} + \Delta u_{it} \end{aligned}$$

(2):  $fd\_afford\_det_{it}$

$$\begin{aligned} &= \beta_0 + \beta_1 fd\_births\_det_{it} + \beta_2 fd\_deaths\_lag_{it} + \beta_3 fd\_immig_{it} + \beta_4 fd\_emig_{it} \\ &+ \beta_5 marr_{it} + \beta_6 divor_{it} + \beta_7 fd\_urban\_pop_{it} + \beta_8 fd\_unem\_30to50_{it} \\ &+ \beta_9 fd\_unem\_rest_{it} + \beta_{10} fd\_constr\_wage\_det_{it} \\ &+ \beta_{11} fd\_prop\_prim_{it} + \beta_{12} fd\_prop\_sec_{it} + \beta_{13} fd\_inc\_gdp\_pc_{it} + \Delta v_{it} \end{aligned}$$

for  $\forall i, t: i \in 1; \dots; N, t \in 1; \dots; T$ , where  $\{\Delta u_{it}\}, \{\Delta v_{it}\}_{i=1;\dots;N,t=1;\dots;T}$  are first differenced idiosyncratic errors.

---

<sup>24</sup> For reasons discussed later on, the *prop\_ter* was excluded.



Under a set of restrictive assumptions, FD estimators can be shown to be the best linear unbiased estimators. (1) Often the most problematic of these assumptions is the strict exogeneity. In the data at hand, temporal dimension and cross-sectional dimension of the panel are both equal to 14. An immense number of lagged variables would need to be included in order to assure validity of the named assumption. For the same reason, detrending of some variables was preferred to including a long series of dummy variables for different years.

An alternative approach is to modify the strict set of premises needed for unbiasedness and rely on a set of assumptions which would at least suffice for consistency of the estimators. In order to take advantage of the law of large numbers and the central limit theorem, the individual processes in our model should be integrated of order zero. (Wooldridge 2013) We fix the problem of non-stationarity by first detrending dependent variables which are prone to trend. Upon examination, these variables were found to be *births*, *constr\_wage* and our dependent variable *afford*<sup>25</sup>. After detrending, the processes are subject to four different unit roots tests.

Levin–Lin–Chu, Breitung, Im–Pesaran–Shin and Hadri Lagrange multiplier stationarity test have been applied. All the information provided in the description of tests was procured from (Stata Corp 2009).

As for the LLC, though author-recommended size of the panel begins at 10 in cross sectional, and 25 in temporal dimension, the available detrended data seem to best fit the asymptotic assumptions on which the test is built. As we will see later, the number of observations, or rather the dimensions of the panel is the biggest limitative aspect of this study. Breitung test can be made heteroscedasticity robust, however for this it requires more time periods without lags than panels. It is praised to have good test power with small datasets such as (25, 25). It is based on the assumption that T and N converges to infinity sequentially. Im–Pesaran–Shin unit root test is the first test in the series to relax the assumption of a common autoregressive parameter for all the panels. Even if the results are said to vary greatly depending on the presence or absence of serial correlation, the test can be made serial correlation robust relying on sufficiently large N and T dimensions.

As all results for these three tests were significant at 1% and suggested no presence of unit roots, they were not included in this paper. On the other hand, p-values differed largely for the Hadri Lagrange multiplier stationarity test and are therefore displayed in a table below:

Hadri LM test is different from the remaining tests in the proposed hypotheses – it tests the null of no stationarity of panels against presence of unit roots in some of the panels. The test can be made cross-

---

<sup>25</sup> To visualize these trends, tables depicting evolution of these variables in time are available in appendix

sectionally heteroscedastic as well as serial-correlation robust and was proposed for panels where the cross sectional dimension is smaller than the time dimension.

Tested process →	fd_afford_d et	fd_nat_inc done	fd_births_d et	fd_deaths	fd_net_mig
p-values →	0.917	0.133	0.713	0.958	0.962
fd_immig	fd_emig	fd_marr	fd_divor	fd_urban_p op	fd_unem
0.09	0.138	0.971	0.966	0.953	0.548
fd_unem_30to5 0	fd_unem_re st	fd_constr_wage_ det	fd_prop_pri m	fd_prop_sec	fd_gdp_pc_grow th
0.739	0.984	0.554	0.973	0.971	0.989

As can be seen from the results, after first-differencing our variables, all are integrated of order zero. Now the issues of serial correlation and homoscedasticity have to be addressed in order for the usual inference of estimators to be hold.

Tests for homoscedasticity are performed based on the assumption of no serial correlation among idiosyncratic errors. (Wooldridge 2013)) Such fact effectively establishes the order in which the tests should be carried out – even though tests for serial correlation require homoscedastic errors, the lack of this property can be compensated by using heteroskedasticity-robust statistics<sup>26</sup>

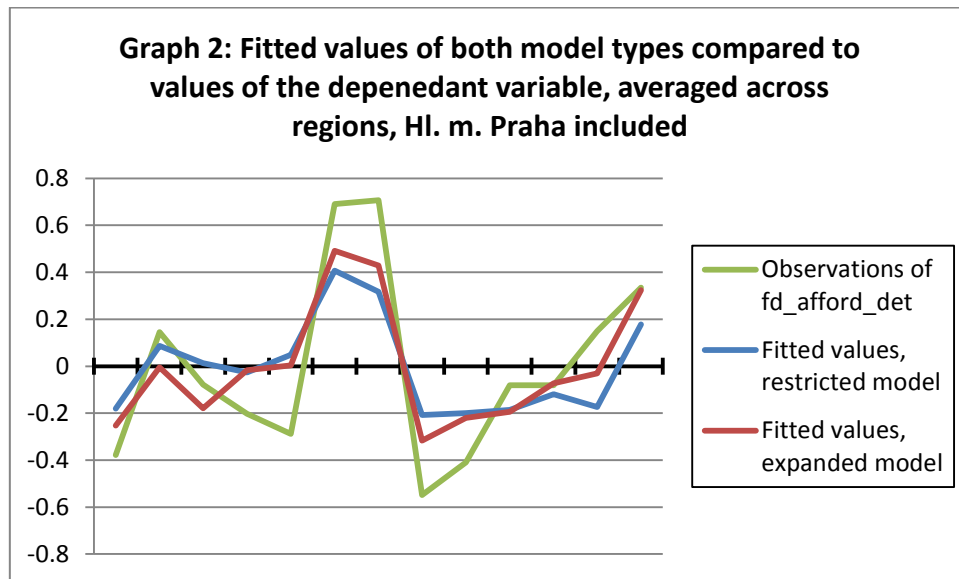
First-differencing might have helped with autoregression of the errors of the first order. Nevertheless, based on author's calculations<sup>26</sup>, higher orders of autoregression among idiosyncratic errors are found. To remedy the defaults in the form of serial correlation and heteroscedasticity of errors, robust t and F statistic will be used in computations.

#### **d. Results of the estimation:**

Before discussing concrete results, it needs to be noted that their validity - that is to say consistency of estimators and their inference -relies heavily on their asymptotic properties. Since the size of the panel both in temporal and cross-sectional dimension is not immense and in comparison with panel data used frequently is rather small, the results should be taken with a pinch of salt.

<sup>26</sup>In the heteroscedasticity-robust frame for testing AR(q) serial correlation proposed in (Wooldridge 2012), second to eighth orders of autoregression in idiosyncratic errors have been detected with results significant at 1% significance level.

Explanatory power of included models measured by adjusted  $R^2$  ranges from approximately 25% to 36% depending on model and data specifications. In this regard, better results are achieved by expanded models and the datasets containing the capital.



An analysis of variability of fitted values, similar to the one performed in the previous subchapter on all the employed determinants, where actual values of  $fd\_afford\_det_{it}$  were taken as means and were compared to fitted values, reveals that determinants applied determinants explain the housing affordability in the case of Prague sufficiently well. There are four regions which performed worse and the overall variability of fitted values for Prague is below the average<sup>27</sup>.

As for the positive or negative impact of individual factors, the majority of the results can be explained by intuition and stylized facts provided in the chapter dedicated to determinants. The results are included as an appendix to this thesis.

Natural increase in population is found to have a negative impact on overall housing affordability. It is worthwhile to note that the measure does not reflect the presumably correct temporal impact of deaths and therefore should be avoided in use, if possible. Without lag, *deaths* variable though statistically insignificant is repeatedly found to have a negative impact on housing affordability. This sign changes instantly once the lag is incorporated. After partitioning  $fd\_nat\_inc$ , the expected negative relation between the relative number of births and affordability is found. The explanation holds that new-borns are believed to contribute to the need of housing whilst at the same time they present another head with no income of its own. Rising relative number of deceased was also expected to have a positive impact on regional housing affordability, once enough time is allowed to pass after death arrived.

<sup>27</sup> Computations done by author not included.

Net migration is a factor negatively affecting housing affordability. Effects of partial components turned out exactly to expectations. With immigrants, the apparent need for housing and the consequential increase in housing prices seems to dominate over the potential increase in disposable income. The statistical insignificance of emigration, once Prague was excluded from the sample, was found to be caused by inclusion of unemployment as an independent variable which seems to suggest<sup>28</sup> that the propensity to migrate is higher among the unemployed.

The one result which escapes readily available explanations is the estimated positive impact of newly formed marriages on affordability. It is possible that the need for new housing is not so strong shortly before and after wedding. Possibly, since the relationship between partners gains official status, those who lived separately move in together and hence stimulate the supply of rather than demand for housing. One can also find a body of literature which deals with wage premium for married men<sup>29</sup>. Whether this premium can offset the debated economic disadvantages marriage is found to have for women is dubious. Author believes that in the absence of firm explanation, the statistically significant determinant should be disregarded.

A rise in number of divorces per 1000 inhabitants exhibits a negative influence on housing affordability. As already discussed, a rising number of divorces is believed to push both on the housing demand as well as reduce the disposable income due to consequences of matrimonial division of labour. Statistical significance of the estimator can be seen as a manifestation of such twofold pronounced effect on both of the parts of housing affordability.

The variable *fd\_urban\_pop* has a positive, at 5 % significance level statistically significant effect on housing affordability which resists invalidation through dataset or model specifications.

In labour market measures, unemployment rate is found to have a positive, highly statistically significant on housing affordability. This suggests that the loss of employment can be compensated by state aid but that it affects greatly the decision of a household whether to acquire an apartment. Separation into age groups proves to be crucial. Largely different wage premium connected with the two proposed age groups is a decisive element in deciding whether unemployment rate will affect the affordability positively or negatively, when the measure is partitioned. Although not sensitive to inclusion of Prague, the *unem\_30to50* proves to be highly statistically insignificant.

After running a series of tests with models which all included detrended version of *prop\_uni* variable, mixed results as to the sign of the effect on affordability were being reported. These tests were conducted across samples with and without the capital and suggested one thing in common for the

---

<sup>28</sup> A suggestion in line with the literature on migration provided in the determinants chapter and further testing done by author.

<sup>29</sup> Wage premium for men in US

prop\_uni variable and that was its high statistical insignificance<sup>30</sup>. For these reasons, the variable was abandoned and was excluded from the final estimations, results of which are incorporated in this study.

More interesting results surfaced with groups of people who have attained either secondary, none or only primary education. Strong economic and statistical significance of the variable capturing the proportion of population with highest attained primary education or no education dictates that amidst potential correlation with other variables it should be regarded as one of the key determinants of housing affordability. The same cannot be said about similar measure of proportion for the people with secondary education. Even though the results are not sensitive to inclusion or exclusion of Prague from the sample, they are different for the restricted and expanded models suggesting some kind correlation with other included variables, probably unemployment. The estimators are also highly statistically insignificant.

In all of the estimations done by author of which only four final are displayed, detrended average gross wage in construction proved to have a strongly statistically significant negative effect on housing affordability. For this reason and for the very plausible economical explanation of influence of the variable on housing prices, it can be established that average gross wage in construction indeed is a housing affordability determinant.

Testing on the first differences of *inc\_gdp\_pc* revealed a positive effect on housing affordability. However, all the encountered results were strongly statistically insignificant. The variable was proposed instead of simple regional GDP per capita, since the latter was found to contained unit roots even after first differencing. It is viewed as a mere proxy and therefore author rejects to associate statistical insignificance of this variable with the same insignificance of GDP per capita in explaining housing affordability and rather stands by insignificance induced by information loss.

When comparing explanatory power of the model via the measure of  $R^2$  and observed first differences, both the expanded and reduced models performed better when Prague was included. Statistical significance of the results viewed at 10% level does not appear to be notably touched by inclusion or exclusion of the capital<sup>31</sup>. Although the absolute size of coefficient slightly differs for regression with and without Prague, for all the statistically significant variables there is no difference in their sign. Very similar conclusions could be drawn from (Hlaváček & Komárek 2009), who used fixed effects rather than first differences in studying housing prices or from a recent thesis by (Cempírek 2014) which uses error-correction model. So far it seems that regional estimations of housing prices

---

<sup>30</sup> Even at significance levels of 70% and above.

<sup>31</sup> Disregarding the measures of emigration, natural increase - latter one is very close to being significant once Prague is excluded from the sample.

determinants and housing affordability can be performed on regional data including Prague without invalidation of the results for the rest of Czech regions.

In comparison with (Hlaváček & Komárek 2009), the impact of net migration, unemployment rate and natural population growth on housing prices is essentially the same as their impact on housing affordability<sup>32</sup>. The results also point to the need to partition these indicators when possible. Not doing so means including variables which can be statistically insignificant<sup>33</sup> and hence invalidate the results for the reduced indicator. In contrast to the named study, divorce rate as well as other newly added variables, namely *fd\_urban\_pop* and *fd\_prop\_prim*, turned out to be statistically significant.

Based on estimations, of which some are provided in this study, and the general discussion about affordability determinants, relative quantity of births, divorces, immigrants and number of people with highest attained primary or no education, along with average gross wage in construction can be classified as housing affordability determinants.

## **V Conclusions:**

The thesis you are currently reading was intended to provide an insight on determinants of housing affordability. As a measure, price-to-income ratio on apartment prices was selected and calculated. Various determinants were included into the proposed model on the assumption that they were affecting both the housing prices and the household income or they were affected only one of the two but were highly correlated with other relevant determinants. The rationale behind this selection was to eliminate those determinants, for which their usual attributed effect either on housing prices or household income wouldn't be touched by their introduction into the model on housing affordability. In this way, avoiding unnecessary repetition of established results can be achieved.

Once the model was constructed, estimation of first differences OLS followed. Among the statistically and economically relevant determinants were the relative numbers of births and divorces, as well as proportion of urban population. The estimation also provided answers the secondary research questions. The results appear robust to the exclusion or inclusion of the Hl. m. Praha region. Prague in itself seems to follow the determinants sufficiently well.

The results, both in their consistency and inference, and hence conclusions based on them however depend to a great extent on the validity of asymptotic OLS assumptions. As a potential extension of this study, a similar model could be estimated on the EU-SILC microdata. This could provide enough

---

<sup>32</sup> Only difference is the insignificance of natural population growth for the estimation without Prague

<sup>33</sup> Such was the case for *fd\_emig* and *unem\_30to50*. Unlagged dependent variable *fd\_deaths* was not only highly statistically insignificant but also going against any explanation and probably contributed to the overall insignificance of natural population growth.

confidence in the results to overstep the bounds of positive empirically estimated determinants on housing affordability and move into the realm normative policy suggestions.

## VII Bibliography

- ABRAHAM & HALTIWANGER (1995): "Real Wages and the Business Cycle", *Journal of Economic Literature*, Vol. XXXIII, pp. 1215-1264
- BAKALOVÁ, J., BOHÁČEK R. MÜNICH D. (2015): "A comparative study of retirement age in the Czech Republic. Prague", IDEA at CERGE-EI, Study (IDEA and SHARE at CERGE-EI)
- BIČÁKOVÁ, A., KALÍŠKOVÁ K. (2015): "Od mateřství k nezaměstnanosti: postavení žen s malými dětmi na trhu práce." *Národohospodářský ústav AV ČR*, v.v.i., Studie Institutu pro demokracii a ekonomickou analýzu.
- BLANCHFLOWER, D. G. & A. J. OSWALD (1989): "The wage curve." *Technical report, National Bureau of Economic Research*.
- BLANCHFLOWER, D. G. & A. J. OSWALD (2005): "The wage curve reloaded." *Technical report, National Bureau of Economic Research*.
- CAHILL, I. G. & M. P. GAGER (2014): "Explaining canadian regional wage differentials." *Review of Regional Studies* 44(2).
- CARD, D. (1999): "The causal effect of education on earnings." *Handbook of labor economics* 3: pp. 1801-1863.
- Cempirek P. (2014): "Housing Prices in the Czech Republic and Slovakia: Regional Comparison", Bachelor Thesis, Charles University of Prague, Faculty of Social Sciences, *Institute of Economic Studies*.
- CIEŚLIK, A., B. ROKICKI ET AL. (2013): "Regional wage determinants in Poland: The empirical verification of the neg approach." *Bank i Kredyt* 44(2): pp. 159-174.
- CSO (2012): „Význam vzdělání pro trh práce v ČR”
- DICKEY, H. (2014): “The impact of migration on regional wage inequality: a semiparametric approach.” *Journal of Regional Science* 54(5): pp. 893–915.
- EARLE, D. (2010): "Skills, qualifications, experience and the distribution of wages." *Tertiary Sector Performance Analysis and Reporting, Strategy and System Performance, Ministry of Education*.
- ÉGERT, B. & D. MIHALJEK (2007): "Determinants of house prices in central and eastern Europe." *Comparative economic studies* 49(3): pp. 367-388.
- FENDEL, T. (2016): "Migration and regional wage disparities in Germany." *Journal of Economics and Statistics*; 236(1): pp. 3–35
- FITZGERALD, J. (2005): "The Irish Housing Stock: Growth in Number of Vacant Dwellings," Quarterly Economic Commentary: Special Articles, *Economic and Social Research Institute (ESRI)*, vol. 2005(1-Spring), pp. 1-22.
- GALUŠČÁK, K. & D. MÜNICH (2002): "Microfoundations of the wage inflation in the Czech Republic." *Czech National Bank, Na příkopě* 28(115)



- HLAVÁČEK, M. & L. KOMÁREK (2009): "Housing price bubbles and their determinants in the Czech Republic and its regions."
- HM TREASURY (2003): "Housing, consumption and emu.-(2003)." *UK Membership of the Single Currency-An Assessment of the Five Economic Tests*.
- HRACHOVEC M. (2012): "Residential Real Estate Market During the Financial Crisis – Empirical Evidence from the CEE Region", Diploma Thesis, Charles University of Prague, *Institute of Economic Studies*.
- KOSTELECKÝ, T. & J. VOBECKÁ (2009): "Housing affordability in Czech regions and demographic behaviour-does housing affordability impact fertility?" *Sociologický časopis/Czech Sociological Review* pp. 1191-1213.
- KUPETS, O. (2016): "Education-job mismatch in Ukraine: Too many people with tertiary education or too many jobs for low-skilled?" *Journal of Comparative Economics* 44(1): pp. 125-147.
- Lasloková L. (2015): "Overeducation in the Czech Labour Market", Bachelor Thesis, Charles University of Prague, Faculty of Social Sciences, *Institute of Economic Studies*.
- LÓPEZ-BAZO, E. & E. MOTELLÓN (2012): "Human capital and regional wage gaps." *Regional Studies* 46(10): pp. 1347-1365.
- LUONG, M. & B.-P. HÉBERT (2009): "Age and earnings." *Perspectives on Labour and Income* 21(1): p. 19.
- MACHIN, S., K. G. SALVANES, & P. PELKONEN (2012): "Education and mobility." *Journal of the European Economic Association* 10(2): pp. 417–450.
- MANKIW, N. GREGORY, DAVID ROMER, AND DAVID N. WEIL (1990): "A contribution to the empirics of economic growth. No. w3541." *National Bureau of Economic Research*.
- MESSINA, J., C. STROZZI, & J. TURUNEN (2009): "Real wages over the business cycle: Oecd evidence from the time and frequency domains." *Journal of Economic Dynamics and Control* 33(6): pp. 1183–1200.
- MYSÍKOVÁ, M. ET AL. (2014): "Educational mismatch in the czech labour market." *National Training Fund*.
- OECD. (2005): Outlook "Chapter iii, recent house price developments: the role of fundamentals."
- ØSTBYE, S. & O. WESTERLUND (2007): "Is migration important for regional convergence? Comparative evidence for norwegian and swedish counties, 1980–2000." *Regional Studies* 41(7): pp. 901–915.
- PISSARIDES, C. A. & J. WADSWORTH (1989): "Unemployment and the interregional mobility of labour." *The Economic Journal* 99(397): pp. 739–755.
- POULIAKAS, K., D. ROBERTS, E. BALAMOU, & D. PSALTOPOULOS (2009): "Modelling the effects of immigration on regional economic performance and the wage distribution: a cge analysis of three EU regions."
- PRITCHETT, L. (2001): "Where has all the education gone?" *The World Bank Economic Review* 15(3): pp. 367–391.

- PSACHAROPOULOS, G. & H. A. PATRINOS (2004): "Returns to investment in education: a further update." *Education economics* 12(2): pp. 111–134.
- PYTLIKOVÁ, M. (2015): "Rozdíly ve výši výdělků ve vztahu k mateřství a dítěti v rodině. Praha: Národohospodářský ústav AV ČR, v.v.i. ", *Studie Institutu pro demokracii a ekonomickou analýzu*.
- STATACORP, L. (2009): "Stata longitudinal data/panel data reference manual release 11."
- SUNEGA, P. & M. MIKESZOVÁ (2015): "Regional disparities in housing affordability in the Czech Republic."
- SUNEGA, P. & M. MIKESZOVÁ, & M. LUX (2009): "Příčiny regionálních nerovností ve finanční dostupnosti bydlení v ČR."
- ŠATAVA J., JANSKÝ P. (2015): "Ekonomické důsledky rozvodu pro oba manžele v České republice: Nákladné rozvody pro ženy či muže?" *Národohospodářský ústav AV ČR, v. v. i. , CERGE-EI, společné pracoviště UK v Praze a NHÚ AV ČR, v. v. i.*
- TROSTEL, P., I. WALKER, & P. WOOLLEY (2002): "Estimates of the economic return to schooling for 28 countries." *Labour economics* 9(1): pp. 1–16.
- WEISS, C. T. (2015): "Education and regional mobility in Europe." *Economics of Education Review* 49: pp. 129–141.
- WOOLDRIDGE, JEFFREY M. (2013): "Introductory econometrics: a modern approach. " 5th ed. Mason, OH: *South-Western Cengage Learning*, xxv, pp. 881
- WRAGE, P. (1981): "The effects of internal migration on regional wage and unemployment disparities in Canada." *Journal of regional science* 21(1): pp. 51–63.

## Appendix

Table 3: First-difference OLS estimation results				
Model type →	Reduced		Expanded	
Capital excluded →	Yes	No	Yes	No
fd_nat_inc	0.1339	0.1077		
fd_births_det			0.3658***	0.3546***
fd_deaths_lag			-0.0039	-0.0025
fd_net_mig	0.0382**	0.5441***		
fd_immig			-0.0222	-0.0539**
fd_emig			0.0369**	0.049***
fd_unem	-0.0838***	-0.083***		
fd_unem_30to50			-0.0801***	-0.0873***
fd_unem_rest			0.0098	0.0262
fd_urban_pop	-0.2169**	-0.2291**	-0.2222***	-0.232***
fd_marr	-0.1178	-0.1806	-0.2351	-0.2677*
fd_divor	0.3148**	0.3131***	0.2971**	0.3138***
fd_constr_wage_det	0.0002***	0.0002***	0.0002***	0.0002***
fd_prop_prim	12.4961*	14.5872**	11.4504**	13.1688**
fd_prop_sec	-0.2563	-0.547	-1.043	-1.5106
fd_gdp_pc_growth	-0.3672	-0.2426	-0.1908	-0.2172
constant	0.0701	0.0829	0.0759*	0.0842*
Adjusted R <sup>2</sup>	0.29051118	0.3206	0.33967887	0.35843492
Results with * are significant at 10% significance level, ** and *** at 5 and 1 percent respectively				