

**NBP**

Narodowy Bank Polski

---

NBP Working Paper No. 182

---

**Papers presented during  
the Narodowy Bank Polski Workshop:  
*Recent trends in the real estate market  
and its analysis, 2013***

Volume 2





NBP Working Paper No. 182

---

**Papers presented during  
the Narodowy Bank Polski Workshop:  
*Recent trends in the real estate market  
and its analysis, 2013***

**Volume 2**

Print:  
NBP Printshop

Published by:  
Narodowy Bank Polski  
Education & Publishing Department  
ul. Świętokrzyska 11/21  
00-919 Warszawa, Poland  
phone +48 22 653 23 35  
[www.nbp.pl](http://www.nbp.pl)

ISSN 2084-624X

© Copyright Narodowy Bank Polski, 2014

---

# Contents

---

11. Hanna Augustyniak, Jacek Łaszek, Krzysztof Olszewski, Joanna Waszczuk: Housing market cycles – a disequilibrium model and its application to the primary housing market in Warsaw .....	5
12. Robert Leszczyński, Krzysztof Olszewski: Panel analysis of home prices in the primary and secondary market in 17 largest cities in Poland .....	39
13. Franz Fuerst, Wayne T. Lim, George A. Matysiak: Non-listed real estate funds: leverage and macroeconomic effects .....	73
14. Wojciech Doliński: The development of commercial property prices and values analysis .....	97
15. M.J. Bárcena, P. Menéndez, M.B. Palacios, F. Tusell: Measuring the effect of the real estate bubble: a house price index for Bilbao ...	127
16. Branimir Jovanovic: Recent developments in house prices in Macedonia .....	157
17. Martin Lux, Petr Sunega: Impact of the economic crisis on house prices in the Czech Republic measured on hedonic price index on bank data .....	173
18. Marta Widłak: The hedonic housing price index for Poland – modelling on NBP BaRN data ...	203
19. Gyula Nagy: Monitoring 15 years of residential house price development in Hungary with the help of the FHB House Price Index .....	231
20. Łukasz Mach: Logit modelling as a tool supporting decision making in the real estate market .....	253

---



---

## **11. Housing market cycles – a disequilibrium model and its application to the primary housing market in Warsaw**

**Hanna Augustyniak, Jacek Łaszek, Krzysztof Olszewski, Joanna Waszczuk**



Hanna Augustyniak – Financial Stability Department, Narodowy Bank Polski.

Jacek Łaszek – Warsaw School of Economics and Financial Stability Department, Narodowy Bank Polski.

Krzysztof Olszewski – Financial Stability Department, Narodowy Bank Polski. ul. Świętokrzyska 11/21, 00-919 Warszawa, Poland, krzysztof.olszewski@nbp.pl. Corresponding author.

Joanna Waszczuk – Financial Stability Department, Narodowy Bank Polski.

We would like to thank an anonymous referee, the participants at the Warsaw International Economic Meeting 2012, the Christmas Financial Stability Seminar 2012 organized by the Bank of Slovakia and the Recent trends in the real estate market and its analysis workshop 2013 organized by the Narodowy Bank Polski for their comments. This paper appeared in its preliminary version in the NBP (2013) “Report on the situation in the Polish residential and commercial real estate market in 2012”.

The paper presents the personal opinions of the authors and does not necessarily reflect the official position of the Narodowy Bank Polski or the Warsaw School of Economics.

---

## Article 11 Contents

1. Introduction .....	9
2. Housing demand .....	13
2.1 Calibration of the demand model and analysis of the impact of interest rates on demand .....	18
3. Supply of real estate developer housing .....	22
3.1 Real and virtual supply curve .....	24
4. Aggregate demand and supply estimation .....	28
5. Summary .....	34
Literature .....	35

## Abstract

This paper presents a simple disequilibrium model in the primary housing market, calibrated to the Warsaw market. Our aim is to point out that the primary housing market, due to the long construction process is always in disequilibrium, which has important policy implications. We discuss the last housing cycle and show how a combination of slight demand shocks with short-term rigid supply leads to strong fluctuations of house prices and new construction. The primary market can create a significant distress to the economy, because when house prices rise, this sector attracts capital and workers and is able to generate excessive supply, which finally can lead to the burst of the price bubble. The cyclical character is a permanent feature of the property market and can be explained by the inelasticity of supply. Market participants form price and demand expectations based on past observations. This causes frequent cycles that, under specific conditions, can lead to economic crises. We believe that the model describes the reality of the primary housing market better than equilibrium models do, so it can be useful for central banks and financial supervision institutions in the analysis of the impact of fiscal and monetary policy and regulations on the real estate market.

**Key words:** Cycles in the housing market, disequilibrium, imbalances, banking sector, banking regulations.

**JEL Classification:** E32, E44, E37, R21, R31.

## 1. Introduction

While modelling the real estate market one usually assumes it to be in equilibrium. However, as a result of a delayed response of supply and changing demand, largely determined by loan availability, the market oscillates around the equilibrium point, fluctuating in time. In this paper we propose a model that accounts for housing needs, which are reflected in a fluctuating housing demand. Supply from the primary market (new construction), which is constant in the short-term, responds with price increases first. Only after a period of time, more new homes are built but prices usually keep growing. Our model allows us to analyze the response of the primary housing market to changes in nominal interest rates or household preferences. We neglect the existing housing stock in this paper and refer to Augustyniak et al. (2014), where we describe the interactions between the primary and secondary housing market in detail. The existing housing stock market hardly faces excessive supply, which is only possible if there is strong migration or if there is very much cheap supply in the primary market. Usually the existing housing stock is owner occupied and changes of its value have a relatively mild effect on the economy<sup>1</sup>. But if we look into the primary market, we observe, for example in the US or Spain, that new construction has created serious problems to the economy and financial stability. New construction is prone to boom and bust cycles, and therefore we analyse the demand and supply of new housing and its cycle.

---

<sup>1</sup> Positive effects of housing appreciation on consumption are well known in the literature. On the other hand if the housing value falls below the loan amount, the owner will not be willing to sell its house, which might impede his mobility.

The analysis of consumer goods requiring a long construction process was described in 1928 by Hanau on the example of a "hog cycle". The investment process and construction of new housing was described, *inter alia*, by Topel and Rosen (1988) and Augustyniak et al. (2012). Our observations indicate that households purchase not only completed housing, but also real estate development contracts for home construction. When demand is largely unmet, real estate developers sell contracts to provide housing, whose construction has already begun. Typically, contracts are sold after approximately one year from the construction start date. Another two years elapse before the apartment is completed. However, after the housing has been constructed, it ceases to affect the housing market, as the contract had already been sold in the past. It should be noted that real estate developers have some work in progress, which allows them to respond to the market needs relatively quickly<sup>2</sup>. Housing pre-sale contracts are known in the US, Italy, Poland as well as in Asia, where this system has had a long history (see Chang and Ward, 1993). Such solutions increase the support elasticity, shorten the cycle and reduce the amplitude of fluctuations. This helps the supply side to respond faster to strongly rising demand, yet, it involves a certain risk. The advantage for clients is that they buy housing at a fixed price. Yet, the buyer bears the risk of the developer's bankruptcy. The developer can continue construction without the need to borrow funds, because the client pays upfront, but he will not be able to increase home prices in the future, along with rising prices or costs.

---

<sup>2</sup> The number of building permits obtained by real estate developers usually exceeds the number of actually started constructions. Moreover, not all the constructions started are immediately sold. The real estate developer can extend this process when prices fall and speed it up when prices rise.

Figure 1 Commenced housing construction, sold real estate development housing (units, left-hand axis) and real price per square meter (2004 constant prices in PLN, right-hand axis)



Source: CSO, NBP BaRN, PONT Info, REAS.

Our analysis focuses on Warsaw, the largest housing market in Poland, but given that most of the world experienced a similar housing boom, our analysis can be easily applied to any other market. The available data allow us to capture the last cycle in the Warsaw residential market. It began with stable prices (2002-2004), which then followed an upward trend with rising income, growing supply of credit and declining interest rates (2005-2008). In the subsequent period (2009-2013) prices slowly decreased as a result of economic downturn, oversupply of unsold housing and limited supply of credit. The relationship between loan availability, growing housing demand and rising home prices in the primary market in Poland is discussed in detail in the NBP (2011, 2012a, 2012b) reports. Looking at real prices, deflated by CPI (2004 is the base year), it can be seen that the actual rise in housing transactions and real estate development contracts, in response to growing demand, led also to a rise in transaction prices (see Figure 1).

When analysing real estate market cycles we need to use a model that replicates the reality as good as possible. One can find very complex models in the literature that describe the housing market. However, in order to solve them, their authors make simplifications and make assumptions which are inconsistent with the reality. In particular, the assumption of the market's rapid search for equilibrium, the assumption about a perfect competition in the market as well as rational behaviour of buyers and producers are problematic. Those simplifications often lead to trivial or even erroneous conclusions. The model presented by us is rather simple and can be replicated on a spread sheet. We show that a fairly simple method allows to analyze imbalances and cycles in the housing market. This requires, of course, relevant data, which, for most of the analysed period, are publicly available on the NBP and CSO website. In our model we go back to the tradition started by Fair (1972) and DiPasquale and Wheaton (1992), who explain the working of the market in a rather simple way.

Our article presents the dis-equilibrium model, analyses cycles and the impact of shocks on the housing market. Chapter 2 presents the model of housing demand. In Chapter 3 we present the model of supply. Chapter 4 presents a regression of aggregate demand and supply as well as price and construction cost changes and Chapter 5 concludes the article.

## 2. Housing demand

In this section we present a simple demand model. While the housing market consists of new construction and sales from the existing stock, it is well known that only the supply of new construction can adjust to a credit fuelled housing demand boom in the short and medium run (Augustyniak et al. 2014). The supply from the existing housing stock is rigid in the short and medium run, thus any excessive demand translates very quickly into excessive demand for new construction. Therefore, we focus on the primary market only and we assume that households finance home purchase with a loan. The cost incurred in a particular period by the household is the loan repayment. Burnham (1972) quotes a Fed survey, according to which credit supply determines housing construction. Currently, we see that housing demand both in Poland and across the world is driven by credit supply (see NBP, 2012a,b). Moreover, demand is affected by consumer preferences as regards the consumption of other goods  $C$  and housing services. Like Bajari et al. (2013) we include the imputed rent  $k * p * H$  in the utility function. It results from the size of the apartment  $H$ , its price  $p$  and the parameter  $k$ , which reflects the monetary value of the stream of housing services. The utility is described by a CES function, where  $\theta$  is the share of utility resulting from consumption, whereas the parameter  $\mu$  denotes the elasticity of substitution between consumption and housing,  $\varepsilon = 1 / (1-\mu)$ . Accounting for appreciation,  $A = \frac{p_t}{p_{t-1}}$ , we take into consideration consumer expectations about future housing prices (see Dunskey and Follain, 1997, Sommervoll et al., 2010 or Lambertini et al., 2012). Such a specification of the utility function takes into account the fact that housing is bought for consumption as well as for investment purposes (see Henderson and Ioannides, 1983 and Łaszek,

2013). It also reflects the fact that households extrapolate past prices and are prone to herding behaviour (see Salzman and Zwinkels, 2013). The utility of household is described by the following equation:

$$U(C, H) = (\theta C^\mu + (1 - \theta)A^\gamma (kpH)^\mu)^{\frac{1}{\mu}}$$

The consumer divides his income between the loan repayment and consumption of other goods. It is very important to remember that not only house prices but also income rises in the real world. If both increase at the same pace, the consumption of housing, measured in square meters should not change. Moreover, after a housing boom relative house prices tend to decline, as the analysis in NBP (2013) shows.

The income allocation problem is solved by taking into account the budget constraint  $b = rpH + C$  ( $r$  – loan constant based on fixed loan instalments,  $p$  – price per square meter of housing), which gives us the optimal choice of the size of housing and consumption of other goods in each period.

$$\theta C^{\mu-1} rp = (1 - \theta)A^\gamma (kp)^\mu H^{\mu-1}$$

By combining this equation with the budget constraint we get the optimum allocation of funds between consumption of housing services and consumption of other goods.

$$C^* = \frac{b}{1 + rp \left( \frac{\theta}{1 - \theta} \frac{rp}{A^\gamma (kp)^\mu} \right)^{\frac{1}{\mu-1}}}$$

$$H^* = \frac{b}{rp + \left( \frac{1 - \theta}{\theta} \frac{A^\gamma (kp)^\mu}{rp} \right)^{\frac{1}{\mu-1}}}$$

Lin and Lin (1999) argue that income elasticity of demand is approx. 1, consequently, income growth should lead to a similar increase in housing demand. The household is not limited by the budget only, but also by the loan availability and supervisory regulations. The bank calculates the loan availability based on household's income, nominal interest rate and prudential regulations, which determine, among other things, the longest possible period of loan repayment (the longer the maturity of the loan, the lower the loan constant and the higher the household loan). When buying an apartment, households usually look at the current market situation and credit granting criteria, neglecting potential changes in interest rates or fluctuations in exchange rates, if they had taken out a foreign currency denominated loan.

In order to curb excessive debt and reduce the risk to the financial system, restrictions are imposed concerning the part of consumer's income which can be spend on the debt repayment (DTI – debt to income ratio) as well as restrictions on the loan to value ratio (LTV) or the down-payment (see Stein 1995, Rubaszek, 2012). For the ease of the analysis, we have considered DTI limits only, so that the household can devote only a part of their income  $x \in (0,1)$  to repay the loan, and the monthly instalment payment is maximum  $b_H$ , whereas:

$$b_H = xb \leq b$$

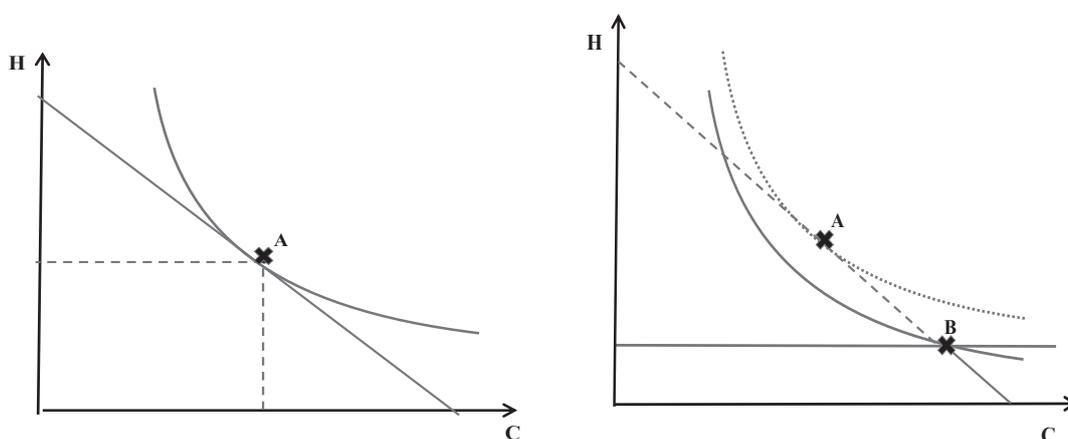
In this situation, the choice of the size of housing will not always be the optimum allocation of funds between  $C^*$  and  $H^*$ , but:

$$H = \begin{cases} H^*, & rpH^* \leq xb \\ \frac{xb}{rp}, & rpH^* > xb \end{cases}$$

$$C = \begin{cases} C^*, & rpH^* \leq xb \\ (1-x)b, & rpH^* > xb \end{cases}$$

Prudential limitations may result in the household’s inability to consume a sufficiently large apartment. On the contrary, the household will be forced to consume more other goods than it needs. Figure 2 shows the consumers choice in the case of a normal budget (point A) and a budget limited by prudential restrictions (point B).

Figure 2 Consumer’s choice without budget constraints (left-hand panel) and with budget constraints (right-hand panel)

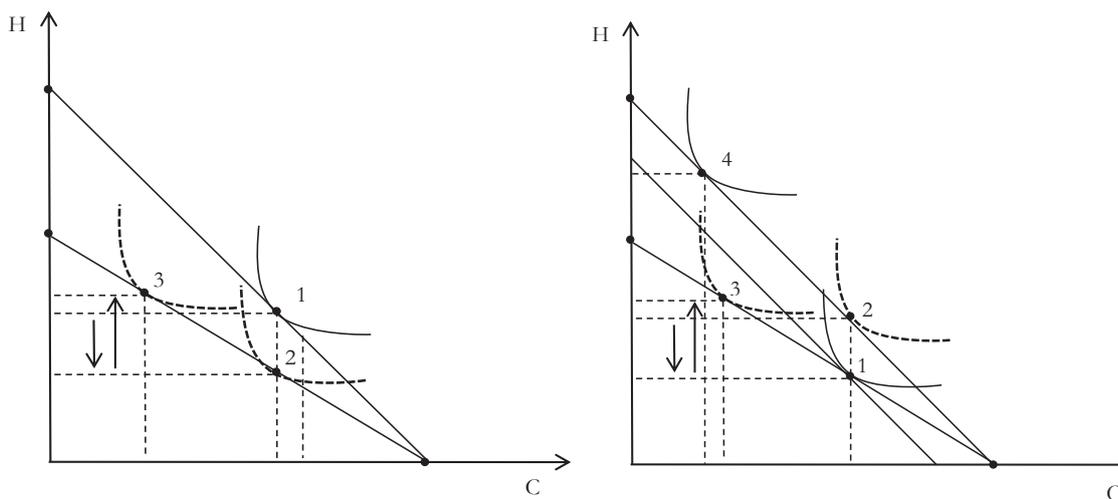


If prudential requirements are restrictive (the household would like to borrow a higher loan than it can), even a slight easing may trigger strong fluctuations in housing demand. Typically, the household is willing to give up some consumption of other goods only to buy more housing. This explains why the availability of foreign currency denominated loans caused a boom in housing demand in Poland.

In order to provide a more thorough explanation of the price bubble, we should present the response of demand to price increases based on a graphical analysis. With the classical utility function, when housing is considered as consumption only, rising prices would lead to a decline in

residential consumption from point 1 to point 2 (see Figure 3, left-hand panel). However, as housing is seen as both a consumption good and an investment good, the housing appreciation causes a significant shift of the utility curve. As a result, amidst rising prices housing becomes even a more desirable asset and the buyer will choose the allocation described in point 3 of the left-hand panel of Figure 3. He will decide to sacrifice even a significant part of consumption of other goods, to buy more housing than he would buy at the former, lower price.

Figure 3 Consumer's choice amidst higher prices (left-hand panel) and an additional increase in income and moreover, interest rate cuts (right-hand panel)



It should be added that prices increased amidst rising incomes and considerable cuts in interest rates. The right-hand panel of figure 3 shows that, in the first place, as a result of rising income the budget line shifts to the right from point 1 to point 2, thus, the consumer can buy more housing and other goods. Yet, prices rise, so as previously explained, he will choose point

3. However, a significant decline in interest rates means that the price increase is, in budgetary terms, almost entirely offset, so the budget curve returns to its position before the price increase, and the buyer finally chooses point 4. The analysis presented in NBP (2013) shows that during the price boom, the loan availability calculated per square meter of housing was on the rise which allowed households to purchase increasingly bigger housing. Only a combination of changes in prices, income and interest rates makes it possible to explain a seemingly irrational behaviour of buyers who amidst rising prices expressed demand for increasingly bigger housing. In the aggregate, this translated into a growing demand for housing.

It is worth noting that in terms of the credit boom, households are able to exceed their budget by taking Ponzi loans. From the perspective of a household's individual decision this can be reasonable, as it optimizes its inter-temporal consumption and will repay the excessive debt in the subsequent period through capital gains on housing. This situation was not observed in the Polish market, thus we do not analyse this variant.

## **2.1 Calibration of the demand model and analysis of the impact of interest rates on demand**

The housing demand model was calibrated to reflect the actual rise in demand and prices in the Warsaw housing market in the years 2002-2013. When calibrating the model presented in the previous section, we have chosen parameters  $\theta$ ,  $\mu$ ,  $k$  that are very close to the values proposed by Bajari et al. (2013) for the US ( $\theta = 0.77$ ,  $\mu = 1.32$ ,  $k = 0.075$ ). We calibrated the model in such a way that it fits the Warsaw housing market as good as possible. We used the same parameter to calculate the substitution elasticity  $\mu = 1.32$ , but

modified the importance of consumption in the utility function to  $\theta = 0.63$ . The smaller parameter  $\theta$  reflects a stronger desire to own housing. The parameter  $k$ , necessary to calculate the imputed rent, was calculated as the average of the actual rental data and transaction prices (from the NBP BaRN database) and amounts to 0.065. Moreover, we choose the parameter, which determines the strength with which buyers react to housing appreciation as  $\gamma = 0.5$ . As the household budget we take the twofold value of the average net wage in the enterprise sector in a particular year. In total, the model replicates the last demand cycle in the primary housing market in Warsaw well, as shown in Figure 4.

The demand for square meters of housing was calculated as follows: the model demand for one square meter of housing of an average household is multiplied by 13 000 - the average number of homes sold during the analysed period, and finally scaled down (by 10 000) in order to harmonize the scale. To calculate the total area of actually sold housing, the number of apartments sold by real estate developers was multiplied by their average size of 58 square meters, and then the score was scaled down (by 10 000). The loan constant was scaled up by 10.

Along with a strong decline in the weighted interest rate<sup>3</sup>, a rise was observed in housing demand. Growth in the model demand largely exceeded growth in actual transactions, which was driven by two reasons. At the beginning of the boom, real estate developers were unable to generate a sufficient number of contracts for housing construction. They put new contracts on the market with one year's delay. Although demand slowed down in the later phase of the cycle, real estate developers sold a lot of contracts – meeting the needs of clients who had expressed their demand a

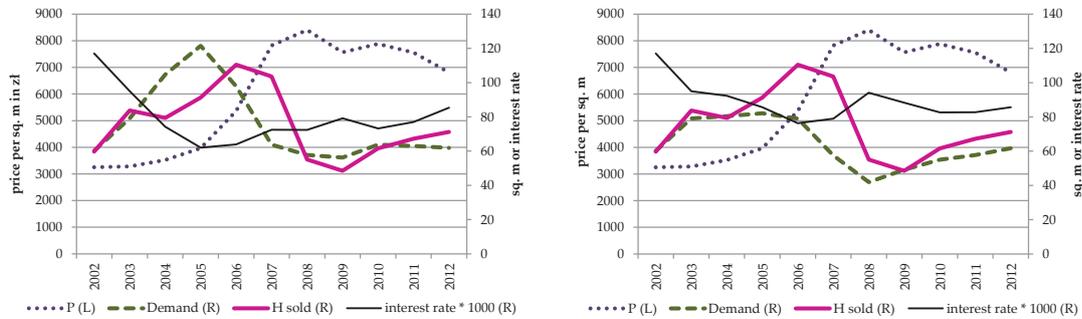
---

<sup>3</sup> The interest rate is weighted with the currency structure of housing loans.

year ago. Moreover, growing income and declining interest rates failed to directly translate into bank lending, which began to pick up with some delay.

The demand model allows us to analyse household behaviour in the primary housing market in the boom period. Although the price of a square meter of housing began to increase rapidly, demand continued to grow. This was driven by three major factors: falling interest rates (due to lower FX interest rates), growing income and expectations of further strong appreciation of transaction prices. This overlaps with the desire to own housing (see Augustyniak et al. 2014) and banks' very lenient lending criteria. A longer crediting period, despite the price growth and a slight increase in the average weighted interest rate was a factor behind rising demand. In this way, the loan instalment remained at a low level for a relatively long time. However, if banks had refused to lend for longer periods much earlier, demand would have probably been lower. Figure 4 shows the development in housing demand, under the assumption that both foreign currency denominated and PLN mortgage loans were granted (left-hand panel) or under the assumption that only PLN mortgage loans were granted (right-hand panel).

Figure 4 Housing demand, home prices, interest rates and the number of purchased housing units (left-hand panel- under the assumption that both PLN and foreign currency denominated mortgages were granted, right-hand panel- under the assumption that only PLN mortgages were granted)



Source: CSO, NBP BaRN, PONT Info, REAS.

The presented model is a partial equilibrium model in which we assume that prices were fixed by real estate developers, and households chose the size of housing only. In fact, reduced demand, resulting from rising interest rates in the absence of foreign currency mortgages should not lead to such a strong price increases (see Figure 4, right-hand panel). It is worth noting that the panel analysis presented at NBP (2013) suggests that real estate developers were quick to raise prices during the boom, yet reluctant to cut them down during demand slump.

### 3. Supply of real estate developer housing

As discussed in the introduction, we focus on the supply of new housing. Housing supply in the secondary market is rather stable in the short and medium run, as it is the result of the willingness of the owner to sell his existing asset. Contrary, housing developers can increase production in the medium-run excessively, as they buy new development land, increase the number of workers and also can sell, besides completed housing, pre-sale contracts. Rising prices attract capital to this market, but the long construction process makes the investment risky – the investor hopes to sell the house for a price that exceeds the construction costs.

Although the price elasticity of housing was already analysed in 1960 by Muth for the United States, the supply side was given relatively little attention in the literature<sup>4</sup>. One of the more extensive publications on this subject is the article by DiPasquale (1999). While analysing housing supply, we must take into account the situation in the local real estate markets. Stover (1986) pointed out that the aggregation of data from individual, distant cities leads to significant errors in the estimation of the price elasticity of housing supply. A detailed analysis of the real estate

---

<sup>4</sup> The modelling of housing supply causes many problems, including analytical ones. Epple, Gordon and Sieg (2010) estimated the home production function based on the price of land. The main problem was that housing consists of attributes that are difficult to value objectively, separately, qualitatively or quantitatively (e.g. quality of housing). It should also be noted that housing features are the result of complex decisions made by real estate developers and home owners (see DiPasquale, 1999). In the case of Poland, the problem is also to gain access to full and accurate data series.

development sector, as well as the profit and cost accounting may be found in Augustyniak et al. (2012), here we focus on the supply curve only<sup>5</sup>.

The short-term supply of developer housing is rigid, because it usually takes around 4 years from the start to the completion of construction. Supply becomes more flexible in the medium term as real estate developers sell contracts, if this is permitted by the law and approved by market participants. Thus, real estate developers put onto the market contracts for home construction, which has only just begun (see Augustyniak et al., 2012) and the whole project will take two more years to complete. During the price boom, the sale of construction contracts began even earlier and the so-called “holes in the ground” were bought. Then, after a period of 3-4 years, completed apartments were delivered to the buyer.

Basing on the relationship between the cost of production and the supply curve of real estate developers, we know that in the medium term the real estate development sector is able to build more housing units at a higher cost. The cost curve will be located close to the marginal cost curve. According to our observations, the mid-term curve of real estate developers' supply may differ significantly from the cost curve, as developers plan future investment based on current prices. They erroneously underestimate the rise in production costs, driven by growing demand and respond only to nominal changes in home prices (see NBP 2013 for more empirical details).

In the medium term, the capital inflow to the residential construction sector pushes the cost of capital down to the level of the minimum average cost (long-term cost). As a result, the supply curve will become even more

---

<sup>5</sup> The price elasticity of supply is explained theoretically by Levin and Pryce (2009) and estimated for different countries by Phang, Kim and Watcher (2010) and by Sanchez and Johansson (2011).

flexible, as new real estate development companies will enter the market, while the existing ones will increase their production. However, if supply rises too much, the average cost will rise as a result of the negative scale effects (infrastructure, costs of transport, materials, land, etc.).

Moreover, real estate developers often fund their operations using financial leverage, which changes their profitability indicators, as increased production financed in this way offsets the growing unit costs. In some countries, it is possible to finance construction with buyers' pre-payments, enabling developers to save the equivalent of interest they would have to pay on the loan, thus increasing the return on investment. Therefore, due to higher prices, the supply of development contracts may be more flexible in the short-term than suggested by marginal costs.

In the long term, supply of housing will be more flexible thanks to a wider range of possibilities of increasing production. The whole economy will be a subject to structural adjustments aimed at adjusting the supply of housing to meet the society's needs<sup>6</sup>.

### 3.1 Real and virtual supply curve

As in most productive sectors, the supply curve is affected by marginal costs and the price. Firms involved in home construction have generally in place similar, standardized construction methods, so that the

---

<sup>6</sup> It is worth noting that the economy may be subject to erroneous, socially expensive and excessive adjustments made to match new housing supply with demand. We have observed this recently, for example in Spain. Too many factors of production (capital and human resources) were transferred to the real-estate development sector, which generated huge costs: a high vacancy rate and mismatches in the labour market.

---

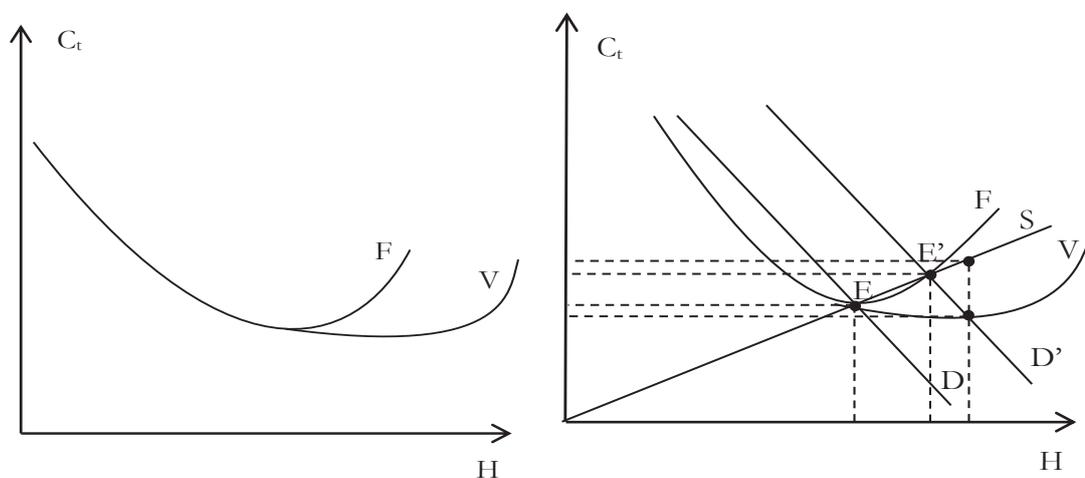
aggregate supply curve is the sum of supply curves of individual real estate developers. We can determine two supply curves: the virtual and the real one.

The virtual supply curve (V) is the result of real estate developer's calculation of future return on investment. This calculation is an estimate based on current housing prices, cost of materials and labour. In contrast to a manufacturing firm, which has a fixed capital stock and an optimal production level above which costs rise substantially, the real estate developer relies on outsourcing of construction services and buys a lot of production factors in small batches. For this reason, the individual cost curve is flat and rises with a considerable delay (see Figure 4, left-hand panel). The real estate developer usually operates as a holding, which allows it to create a special purpose vehicle to start new investment projects.

Furthermore, the number of housing units in a particular location can be adjusted to meet the current market needs. Its supply is limited by the access to capital generated by the stock and bonds market and by loans. We should also mention another restriction in the form of a limited number of qualified people who can conduct the construction process and the availability of production factors. Consequently, real estate developers supply curve (S) will be less inclined. It will move to the left, if the real estate developer expects the cost of land, materials and labour to increase or the diseconomy of scale starts to emerge.

Furthermore, the real estate developer can continue a project that was already started and then stopped, should demand increase. In this case, the supply curve of the developer is virtual and subjective as it is based solely on self-estimates and usually does not take into account the behaviour of the competition.

Figure 5 The virtual and actual supply curve (left-hand panel) and changes in the market in response to rising demand (right-hand panel)



The real supply curve (F) of the entire development sector reflects the actual changes in investment profitability, taking into account diseconomies of scale and increasing costs of production factors, when production reaches too high levels. For example, real estate developers will buy less attractive plots of land and will have to adapt them to meet the actual needs or will have to pay more for work and materials. This gives a curve that shows how flexible the response of demand to housing prices is. However, its importance at the planning stage seems to be limited to the individual developer. It will play a major role in the final phase of the project and will determine the number of housing units that are currently being constructed.

If housing prices are stable, growing costs mean that profit margins are lower and, consequently, the expected future profits decline and the virtual supply curve goes up. This results in the suspension of new investments, slowing construction of the existing real estate investments, or even abandonment of the current projects. The virtual and the actual supply curve is shown in the left-hand panel of Figure 5. Rising demand for

housing, as shown in the right-hand panel of Figure 4, as a shift in the demand curve to the right (from  $D$  to  $D'$ ), will urge real estate developers to increase housing production. As a result of rising factor prices and the negative scale effects (average costs of production factors begin to grow), the real cost of production of such a large number of dwellings is higher (see curve  $F$ ) and exceeds the price that consumers are ready to pay. This, in turn, creates a surplus in the housing market. We would like to mention that housing developers have tools to cover excessive costs, as they apply price discrimination during the selling process. As described in Łaszek and Olszewski (2014), they offer each housing unit to every buyer at his own reservation price, thus obtain a higher profit than they would obtain under perfect competition.

#### 4. Aggregate demand and supply estimation

In this last chapter we investigate empirically the aggregate demand supply and price adjustment process in the primary housing market. Because the annual time series are rather short, we use quarterly data<sup>7</sup>. As the quarterly data shows seasonality and also is more prone to short-term shocks that we do not model here, we use the moving average over four quarters. We base on the previously described model and the housing stock adjustment literature (see Mayer and Sommervoll, 2000 and Steiner, 2010). In aggregate terms there is an equilibrium housing stock level, corresponding to a specific price level. If, for some reason, the demand for the housing stock increases, there is a mismatch between demand and supply and prices rise. As shocks are a permanent feature, the market is never in pure equilibrium, but it may be very close to it. Following Mayer and Sommervoll (2000), we regress the log-levels of housing demand and supply in a given period on price changes and also explain the price and construction changes by the mismatch between demand and supply. Taking into account the previously described model of individual demand for housing space  $H_t$ , we proceed to the aggregate demand  $HD_t$  for housing units. The aggregate demand in log terms can be described by equation 1.

$$\text{Eq. 1. } HD_t = \alpha_1 + \alpha_2 * P_t + \alpha_3 * D(P_t) + \alpha_4 * \text{Intrate}_t + \alpha_5 * \text{Income}_t + \epsilon_t$$

---

<sup>7</sup> The data on house prices origins from the NBP BaRN data base, data on house sales and housing put on the market comes from REAS, while the construction costs are based on Sekocenbud data. The income data comes from the CSO, while data on interest rates comes from NBP. Data on supply, demand, prices and costs as income is in logs.

Here  $P_t$  is the log house price,  $D(P_t)$  is the growth rate of the house price. The interest rate ( $Intrate_t$ ) and the income in log terms ( $Income_t$ ) are included to capture the development of the economy. The empirical analysis (see table 1 below) confirms that in aggregate terms housing demand depends positively on the income, while it declines with price or interest rate increases. Moreover, the housing demand increases when prices are on the rise, which confirms the herding or speculative behaviour of home buyers in the primary market. We want to point out that if price increases at the same time as income, both effects annihilate each other. In relative terms, housing measured in units of income is as expensive as it was before the price increase (see also NBP (2013b) for data on other cities in Poland).

The next step is the regression of the aggregate housing supply in the primary market. Again in line with theory we observe that developers react with a lag of one year to price increases, which is captured by  $D(P_{t-4})$ . If prices rise, developers start new construction projects, which they put with a one year lag on the market. Increasing construction costs, captured by  $D(PC_{t-4})$ , decrease their will to start new projects as well as increasing interest rates (denoted  $Intrate_t$ ) do. The higher the interest rate is, the more costly is a loan or the higher are the opportunity costs of investing in housing construction. Finally, given the high fixed costs, the developer produces a certain amount of housing, irrespective of the current prices and construction costs. We call this production autonomous (see also Augustyniak et al. 2012).

$$\text{Eq. 2. } HS_t = \beta_1 + \beta_2 * D(P_{t-4}) + \beta_3 * D(PC_{t-4}) + \beta_4 * Intrate_t + \epsilon_t$$

The price setting behaviour of the real estate developer is captured by equation 3. First, it turned out that the house price depends strongly on its

past level, thus also the current price growth rate  $D(P_t)$  depends on its past growth rate  $D(P_{t-1})$ . As the literature on housing cycles shows (see for example Bracke, 2013), there are long periods of price growth and even longer periods of price decline. The developer adjusts the price also to the mismatch between current demand and supply (denoted  $HS_{t-1} - HD_{t-1}$ ), that is under excessive demand he increases the price (see Tse Ho and Ganesan, 1999).

$$\text{Eq. 3. } D(P_t) = \vartheta_1 + \vartheta_2 * D(P_{t-1}) + \vartheta_3 * (HS_{t-1} - HD_{t-1}) + \epsilon_t$$

We tested whether there is asymmetry in the price adjustment and found that the price increase in response to excessive demand is as strong as the price decline in response to excessive supply. In the case of oversupply, the developer lowers his price expectations slowly, hoping to find a buyer who is willing to pay the price. This has been observed in the Polish real estate market but also in other international markets in the recent years. Even though a fast price decline would allow the market to clear and reach the equilibrium faster, developers are not willing to lower prices too fast. Typically, developers have price expectations and are ready to wait for a client who is willing to pay their price<sup>8</sup>. Also when the construction is financed with a loan, the loan agreement may prevent the developer from lowering prices below a certain threshold. The buyer can negotiate the price, yet has a small amount of information and little bargaining power<sup>9</sup>. In

---

<sup>8</sup> Compare the offers presented in Figure 4 in NBP (2013).

<sup>9</sup> There is a strong asymmetry of information, the developer can put a smaller number of housing units on the market to create the appearance that housing is a rare good. The

---

consequence, the price remains high and even though there is excessive supply, other developers continue to deliver housing to the market.

The last important factor which affects housing construction are construction costs, which we estimate in equation 4. We observe that construction costs, similar to house prices, tend to depend strongly on their past realization. Moreover, increases in supply require more input goods and rise production costs.

$$\text{Eq. 4. } PC_t = \rho_1 + \rho_2 * D(PC_{t-1}) + \rho_3 * D(HS_{t-1}) + \epsilon_t$$

We estimate each equation separately using the OLS regression, correcting for heteroskedasticity and autocorrelation. The estimation is run on quarterly data, and due to the limited data length the regression 1 covers the period 2004Q2-2013Q4, regression 2 the period 2005Q1-2013Q4, regression 3 the period 2008Q1-2013Q4 and regression 4 the period 2004Q3-2013Q4.

---

housing developer has also marketing tools to convince the client that housing is worth as much as expected by the developer.

Table 1. Regression results of the determinants of aggregate supply, demand, prices and production costs. Each regression estimated separately with OLS.

	HD_t	HS_t	D(P_t)	D(PC_t)
P_t	<b>-0.889 ***</b> (0.180)			
D(P_t)	<b>7.456 ***</b> (1.395)		<b>0.823 ***</b> (0.074)	
D(P_t-4)		<b>9.300 ***</b> (1.841)		
Intrate_t	-8.733 * (4.932)	<b>-16.558 ***</b> (6.105)		
Income_t	0.780 * (0.396)			
D(PC_t-1)				<b>0.894 ***</b> (0.087)
D(PC_t-4)		<b>-13.160 **</b> (2.102)		
D(HS_t-1)				<b>0.023 **</b> (0.009)
HS_t-1 – HD_t-1			<b>-0.033 ***</b> (0.009)	
C	<b>9.828 ***</b> (2.222)	<b>9.062 ***</b> (0.345)	0.001 (0.002)	-0.001 (0.001)
Adj. R <sup>2</sup>	0.82	0.73	0.72	0.92

Newey-West HAC standard errors in brackets, \*\*\*, \*\*, \* significant at the 1%, 5% or 10% level.

The four equations can be used to explain the dynamics in the housing market. We see that persistent low interest rates or increasing wages trigger a demand boom, that is followed by a price increase and supply boom. Due to the appreciation effect and also because with rising income relative house prices remain stable, the housing boom period can last for a quite long period. It will be stopped only by a shock, such as the sub-prime crisis in the USA, which made banks slow down the disbursement of loans.

The model helps to explain the price dynamics and the number of real estate transactions in the primary market. It suggests that the only way to achieve a market equilibrium and ensure small fluctuations around this point, is to stabilize and control demand, among others by slowing down the credit boom. Such a strong demand boom would probably not be possible in Poland, should only zloty denominated loans be granted. Moreover, the regulations related to the existence of the government-subsidized housing scheme sustained the demand (see: NBP, 2013). If, however, a programme aimed to support the rental housing market was introduced, the demand shock caused by rising incomes and low interest rates could be limited. Still, there is the secondary market and it is important to consider this market in the analysis of housing policy. We refer to Augustyniak et al. (2014) and NBP (2013b) for a detailed discussion on housing policies for the secondary housing market.

## 5. Summary

We present a relatively simple model that helps to understand the cyclical nature of the housing market. After calibrating the model to the primary housing market in Warsaw, we show how changes in interest rates affect the demand. Moreover, the model shows how demand shocks affect housing production and prices, and can lead to repeating housing cycles. It may be concluded that only the reduction in demand, for example, by prudential regulations limiting the availability of loan-financed housing can help to smooth out the cycle in the primary housing market.

An important assumption of the model, motivated by empirical observations, is that the primary housing market is in constant disequilibrium. Delayed adjustments of supply to the continuously changing demand lead to permanent cycles. We wish to emphasize that the equilibrium assumption, on which most of the known housing models are based, gives erroneous results and misleading indications for decision makers. We believe that our model is useful for policy makers, central banks and regulators to analyse and understand the impact of various factors in the primary housing market.

## Literature

- Augustyniak, H., K. Gajewski, J. Łaszek and G. Żochowski (2012), *Real estate development enterprises in the Polish market and issues related to its analysis*, MPRA Paper 43347. Appeared also as *Przedsiębiorstwo deweloperskie na rynku w Polsce oraz problemy jego analizy*. Bezpieczny Bank, 2013, 4(53), 227-235.
- Augustyniak, H., J. Łaszek, K. Olszewski and J. Waszczuk (2014), *Modelling of cycles in the residential real estate markets – interactions between the primary and the secondary market and multiplier effects*, *Kwartalnik Nauk o Przedsiębiorstwie* 2/2014 (31).
- Bajari, P., P. Chan, D. Krueger and D. Miller (2013), *A dynamic model of housing demand: estimation and policy implications*, *International Economic Review* 54, 409–442.
- Bracke, P. (2013), *How Long Do Housing Cycles Last? A Duration Analysis for 19 OECD Countries.*, *Journal of Housing Economics*, 22(3), 213-230.
- Burnham, J. B. (1972), *Private Financial Institutions and the Residential Mortgage Cycle, with Particular Reference to the Savings and Loan Industry*, In: Board of Governors of the Federal Reserve System, *Ways To Moderate Fluctuations in Housing Construction*.
- Chang, C.-O. and C.W.R. Ward (1993), *Forward pricing and the housing market: the pre-sales housing system in Taiwan*, *Journal of Property Research* V 10, 217-227.
- DiPasquale, D. and W. C. Wheaton (1992), *The markets for Real Estate Assets and Space: A Conceptual Framework*, *Journal of the American Real Estate and Urban Economics Association*, Vol. 20, 181-197.
- DiPasquale, D. (1999), *Why Don't We Know More About Housing Supply?*, *The Journal of Real Estate Finance and Economics*, Vol. 18, 9-23.

- Dunsky, R.M. and J.R. Follain (1997), *The demand for mortgage debt and the income tax*, Journal of Housing Research, Vol. 8, 155-199.
- Epple, D., B. Gordon and H. Sieg (2010), *A New Approach to Estimating the Production Function for Housing*, The American Economic Review, Vol. 100(3), 905-924.
- Fair, R. C. (1972), *Disequilibrium in housing models*, The Journal of Finance, 27(2), 207-221.
- Hanau, A. (1928), *Die Prognose der Schweinepreise (Forecasting the price of pork)*, Vierteljahreshefte zur Konjunkturforschung Sonderheft 7.
- Henderson, J.V. and Y.M. Ioannides (1983), *A Model of Housing Tenure Choice.*, The American Economic Review, Vol. 73(1), 98-113.
- Lambertini, L., C. Mendicino and M. T. Punzi (2012), *Expectations-Driven Cycles in the Housing Market*, Bank of Finland Research Discussion Paper 2-2012.
- Levin, E. J. and G. Pryce (2009), *What Determines the Responsiveness of Housing Supply? The Role of Real Interest Rates and Cyclical Asymmetries.*, Housing Studies 24(6), 713-736.
- Lin, C. and S. Lin (1999), *An Estimation of Elasticities of Consumption Demand and Investment Demand for Owner-Occupied Housing in Taiwan: A Two-Period Model*, International Real Estate Review, Vol. 2, 110-125.
- Łaszek, J. (2013), *Housing and consumer theory*, In: Report on the situation in the Polish residential and commercial real estate market in 2012, NBP.
- Łaszek J. and K. Olszewski (2014). *Zachowania firmy deweloperskiej na rynku mieszkań i zagregowana podaż*. Mimeo.
- Mayer, C. J., and C. T. Somerville (2000), *Land use regulation and new construction*. Regional Science and Urban Economics 30(6), 639-662.

- Muth, R. (1960), *The demand for non-farm housing*, in A.C. Harberger, *The demand for durable goods*, The University of Chicago Press.
- NBP (2011), *Report on the situation in the Polish residential real estate market in 2010 r.*
- NBP (2012a), *Report on the situation in the Polish residential real estate market in 2011 r.*
- NBP (2012b), *Information on home prices and the situation in the residential and commercial real estate market in Poland in 2012 (Q1 and Q2).*
- NBP (2013a), *Information on home prices and the situation in the residential and commercial real estate market in Poland in 2013 Q1.*
- NBP (2013b), *Report on the situation in the Polish residential real estate market in 2012 r.*
- Phang, S. Y., K. H. Kim and S. Wachter (2010), *Supply Elasticity of Housing*, In: *International Encyclopedia of Housing and Home*, Elsevier and Science Direct.
- Rubaszek, M. (2012), *Mortgage down-payment and welfare in a life-cycle model*, *Bank and Credit*, Vol. 43(4), 5-28.
- Salzman, D.A.; Zwinkels, R.C.J. (2013), *Behavioural Real Estate*, Tinbergen Institute Discussion Paper, No. 13-088/IV/DSF58.
- Sanchez, A. C. and A. Johansson (2011), *The Price Responsiveness of Housing Supply in OECD Countries*, OECD Economics Department Working Papers, No. 837.
- Sommervoll, D. E., T.-A. Borgensen and T. Wennemo (2010), *Endogenous housing market cycles*, *Journal of Banking & Finance*, Vol. 34, 557-567.
- Stein, J. (1995), *Prices and trading volume in the housing market: a model with downpayment constraints*, *Quarterly Journal of Economics*, 110(2), 379-406.

- Stover, M. E. (1986), *The price elasticity of the supply of single-family detached urban housing*, *Journal of Urban Economics*, 20(3), 331-340.
- Steiner, E. (2010), *Estimating a stock-flow model for the Swiss housing market*, Swiss National Bank Working Papers 2010-08.
- Topel, R. and S. Rosen (1988), *Housing Investment in the United States*, *Journal of Political Economy*, Vol. 96, 718-740.
- Tse, R.Y.C., C.W. Ho and S. Ganesan (1999), *Matching housing supply and demand: an empirical study of Hong Kong's market*, *Construction Management and Economics* Vol. 17, 625-633.
- Wheaton, W. C. (1999), *Real Estate "Cycles": Some Fundamentals*, *Real Estate Economics*, Vol. 27,2, 209-230.

---

## **12. Panel analysis of home prices in the primary and secondary market in 17 largest cities in Poland**

**Robert Leszczyński, Krzysztof Olszewski**



Robert Leszczyński – Narodowy Bank Polski, Regional Branch in Białystok.  
Krzysztof Olszewski – Financial Stability Department, Narodowy Bank Polski. ul. Świętokrzyska  
11/21, 00-919 Warszawa, Poland, krzysztof.olszewski@nbp.pl. Corresponding author.

We would like to thank Hanna Augustyniak, Tomasz Chmielewski, Jacek Łaszek and the participants of the NBP Workshop 2013 “Recent trends in the real estate market and its analysis” as well as the SGH “Modelowanie danych panelowych: teoria i praktyka 2014” conference for comments. Also two anonymous referees helped to improve the paper. An initial version of this paper was presented in the NBP (2013) “Report on the situation in the Polish residential and commercial real estate market in 2012”.

The paper presents the personal opinions of the authors and does not necessarily reflect the official position of the Narodowy Bank Polski.

---

## Article 12 Contents

1. Introduction .....	43
2. Empirics .....	48
2.1 Analysis of the primary market .....	52
2.2 Analysis of the secondary market .....	57
2.3 Interdependence between the primary and secondary market and spillover effects ...	61
3. Conclusions .....	67
Literature .....	69

## **Abstract**

We analyse the determinants of house prices in the primary and secondary market of 17 largest cities in Poland during the 2002-2013 period. We find that prices are driven by economic fundamentals, such as income growth or rise in employment. Prices in the secondary market react to increases in the loan availability, that was driven by low interest rates resulting from FX denominated housing loans that were granted since 2006. This finding does not hold for the primary market, which is to a large extent financed with cash. We confirm empirically that the house appreciation in the past period has a strong effect on the current price, which confirms herding behaviour in the housing market. Another finding is that the secondary market has a stronger effect on the primary market than the other way around. This means that housing demand is satisfied in the first place from the secondary market, and if prices rise, potential buyers go to the primary market. Finally, we find that price increases in Warsaw spill over to the local markets of 16 regional cities. This finding is consistent with the contagion theory in the real estate market, according to which price increases in the centre lead to price increases in the periphery.

**Key words:** housing market, house prices, primary and secondary market, spillover effects

**JEL classification:** E21, R21, R31

## 1. Introduction

Growth in home prices in the primary and secondary market is the subject of continuous interest of central banks and regulators, as it rapidly translates into changes in real estate construction, drives housing cycles (see Augustyniak et al., 2014a) and generates risk for the banking sector. The recent international financial crises resulted from a housing boom and bust in the US and the lesson that everybody has learned is that housing is of paramount importance for households and the financial system, thus also for the whole economy. The stiffness of supply in short term can cause price shocks and also create bubbles. Mortgage loans constitute the largest liability of households and also a major asset of banks. Even though only a small fraction of housing is traded at a given time, it affects the price of the whole housing stock. During a price boom households and banks are willing to accept high loans. However, during a price bust the loan value can exceed the collateral, which can cause problems to the stability of the financial system.

Our study focuses on determinants of the average price<sup>1</sup> of a square meter of housing in Poland's 17 largest urban markets in the 2002-2013 period. The housing market is of high importance for the economy, and the 17 largest cities account for around 40% of housing loans issued in Poland. In 2013 around 30% of the total number of housing units (all types of housing) was delivered in this largest markets, and among those completed,

---

<sup>1</sup> Dwellings in the primary and secondary market display differences in terms of building technology, quality of finishing, as well as in location, thus it would be optimal to use the hedonic price, which accounts for the heterogeneity (see Widłak, 2013). However, such an analysis requires very detailed data that have been collected in the NBP BaRN data base since 2006 Q3 only. If we limited the analysis to the period commencing practically in 2007 we would not be able to capture the behaviour of prices during the period of price stability, that is, in the years 2002 - 2005.

nearly 80% were produced by housing developers. Further on, those cities constitute the largest labor market, are the regional capitals of the voivodeships and also a significant part of GDP is created in those cities.

The aim of our analysis is to determine which economic variables explain the growth in home prices in the primary and secondary market. We analyse factors that affect the demand for new and existing housing in the largest cities in Poland, using the findings presented in Andrews (2010) and Igan and Loungani (2012), Cameron et al. (2006) and also the results of the analysis of the convergence and differentiation of regional housing markets and structural changes in Poland presented in Baldowska et al. (2014).

The analysis of the Polish housing market (see NBP, 2013 and Augustyniak et al., 2014b) indicates that there is a high demand for owner occupied housing. The desire to own housing has two main reasons. Renting is quite costly, as in many cases the loan instalment can be even lower than the rent and when the loan is paid back, the owner keeps the house forever. Secondly, there are investment motives, because housing is not only a durable consumer good but also an investment good (see Łaszek, 2013 for a detailed discussion). The desire to own housing can be only satisfied, if the household has enough income to accumulate the down payment and to cover the loan instalments and expects that it will receive income in the future.

The housing demand analysis has a long tradition and we refer to Andrews (2010), Igan and Loungani (2012) and Augustyniak et al. (2014a) for a discussion of theoretical models and empirical results. Factors that increase housing demand are high income and/or low interest rates, that lead to lower loan instalments. Moreover the average unemployment rate is a good proxy for the overall performance of the economy. A falling

unemployment rate indicates that the economy is booming and makes people more optimistic. As concerns the investment demand, the appreciation of housing (change of prices from period to period) makes the buyer to speculate that prices will rise further (see Augustyniak et al., 2014a,c). The long memory of house prices was confirmed empirically by Andre et al. (2013), who performed a cointegration analysis of house prices among countries in the euro area. Andrews (2010) finds in an international analysis of house prices that the preferential tax treatment of house owners adds to the demand shock. The abovementioned factors increase the demand for housing, which under fixed supply in the short run leads to house price growth and we formulate the following research hypothesis.

**Research hypothesis 1:** House prices in the major markets in Poland are determined by fundamental economic determinants and also depend on herding behaviour.

Housing demand can be satisfied with housing from the primary or the secondary market and the consumer choice is determined by price differences and housing availability. The market mechanism, which is described in detail in Augustyniak et al. (2014a), is as follows. First, buyers try to obtain housing from the secondary market, which is usually cheaper. In the secondary market there is a given amount of housing put on the market, for example, because people want to move or because they want to up-size or down-size their housing. If housing demand increases, the supply in the secondary market starts to disappear, yet other households, who did not enter the market so far, become interested in selling their house. Some who had a smaller house decide to sell it and to buy a new, bigger one. At some point the excess demand can be satisfied only from the primary market. But as supply in the primary market is quite rigid in the short run,

prices start to rise. Real estate developers observe rising prices and start to increase the production of new housing and start to sell more pre-sale contracts<sup>2</sup>. The price depends on the construction costs and the profit margin of the developer. We observe interactions between the two markets, investigate them empirically and formulate the second research hypothesis.

**Research hypothesis 2:** There are strong relations between the primary and secondary housing market in Poland.

Recent empirical studies find a contagion effect in the housing market, which means that price increases in one location spill over to other locations. Examples of such studies are Ferreira and Gyourko (2011) for the US, Fry (2009) for the UK and Vansteenkiste and Hiebert (2011) or De Bandt et al. (2010) for the world housing market. Andre et al. (2013) provide a rich overview of the empirical work on the comovement of house prices in different regions. Spillovers can be observed also at the city level, as the recent research for the Warsaw market by Waszczuk (2013) and Widłak et al. (2014) shows. DeFusco et al. (2013) use micro-data for the US and find that the dynamics are stronger during the boom period than during the bust period and the price impulse is stronger from large markets than from smaller markets. Dittmann (2013) analysed this question for Poland and we perform a similar analysis using a longer data set. We formulate the third research hypothesis.

---

<sup>2</sup> Housing pre-sale contracts are common in Poland and Asia (see Chang and Ward, 1993), while they are less common in Western Europe. According to Augustyniak et al. (2014c, p.2) „Such solutions increase the support elasticity, shorten the cycle and reduce the amplitude of fluctuations. This helps the supply side to respond faster to strongly rising demand, yet, it involves certain risk. The advantage for clients is that they buy housing at a fixed price, thereby financing the real estate developer’s project. Consequently, the developer can continue construction without the need to borrow funds. Yet, the buyer bears the risk of the developer’s bankruptcy. On the other hand, the producer of housing will not be able to increase home prices in the future, along with rising prices or costs.”

**Research hypothesis 3:** There are spillover effects from changes in house prices in the capital city to changes in house prices in other major cities.

The remainder of the paper consists of the empirical analysis. In chapter 2 we analyse the primary and secondary market, determine the interactions between those two markets and also search for a spillover effect of price rises in the Warsaw market, that spread to other local markets. Section 3 discusses the main results.

## 2. Empirics

We use the results from the literature, presented in the introduction as a starting point for our empirical analysis. According to the mainstream literature prices of new housing should be affected by structural factors (i.e. the number of new marriages per 1 000 inhabitants, migration, the ratio of the productive-age population to post-productive age population, etc.) as well as economic factors (income growth, unemployment, loan availability). The analysis of transaction prices of housing in the primary and secondary market of 17 cities in Poland is based on annual data for the years 2002-2013<sup>3</sup>. The analysis takes into account the relatively stable period in the housing market (2002-2005), the housing boom period (2006-2008) and the market's slow return to its equilibrium (2009-2013). Due to data limitations this is the longest period that we can analyse, but it captures at least a full real estate cycle and gives a good picture of the determinants of price changes in the residential market. In 16 largest cities the primary market prices are higher than the secondary market prices, while the opposite is true for Warsaw. For presentation purposes we present the CPI deflated house prices in the primary and secondary market in the six biggest cities on graph 1. Usually new construction is of better quality than the existing housing stock, it is more energy efficient and has many amenities, such as parking space under the building or elevators. It also can be directly used without the need of a general refurbishment. However in the case of Warsaw most of new construction took place in parts of the city that are rather distant from the city centre and which initially lacked a well-developed cultural, educational and transport infrastructure, thus potential

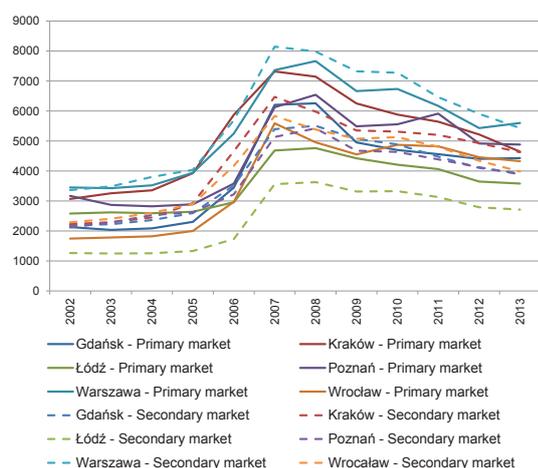
---

<sup>3</sup> Transaction prices of housing for the years 2006-2013 are from the BaRN database (primary and secondary market). Prices in the period 2002-2005 were extrapolated on the basis of primary market offer price growth coming from the PONT Info database .

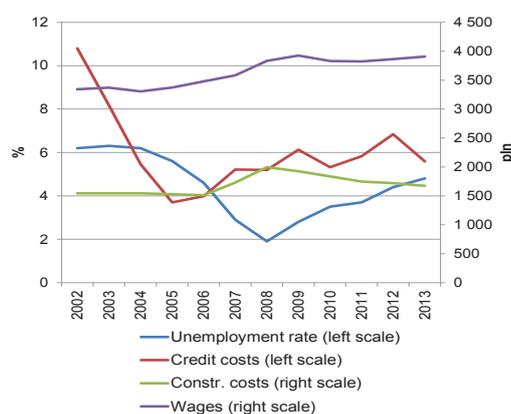
buyers were willing to pay less than for the existing housing stock. According to the estimates presented in NBP (2013), in the regional cities approximately around one third of transactions in the housing market took place in the primary market, while two-thirds took place in the secondary market. This results from the lower prices but also from the much higher supply of housing in the secondary market.

In graph 2 we present the main determinants of house prices for the Warsaw housing market, that is the unemployment rate, credit costs, construction costs and wages. The data shows us that during the whole period wages were rising, while the price boom was preceded and accompanied by a falling unemployment rate and falling credit costs. As a result of the housing boom, the constructions costs started to rise, but declined slowly after the boom faded out after 2008.

Graph 1 House prices on the primary and secondary market – 6 biggest cities (CPI defl.)



Graph 2 Unemployment rate, credit costs, construction costs, wages in Warsaw



The main economic and fundamental drivers of housing demand in the remaining cities behave in a similar fashion. On initial, visual inspection

of the data suggests to use the abovementioned fundamental variables to explain the house price movements in the local markets. Our empirical model draws on the metropolitan area house price model of Jud and Winkler (2002), the abovementioned empirical literature and the detailed analysis and description of the housing market in Poland, presented in Baldowska et al. (2014). We express the demand for housing  $Q_d$  as a function of the following determinants:

$$Q_d = f(\text{price, wages, loan availability, economic conditions})$$

while the housing supply  $Q_s$  is expressed as:

$$Q_s = f(\text{price, construction costs, profits, local determinants})$$

Because demand has to equal supply in equilibrium, we can join the two equations and obtain the determinants of the house price:

$$P = f(\text{wages, loan availability, economic conditions, construction costs, profits, local determinants}).$$

We modify the equation in line with the analytical housing demand model presented in Augustyniak et al. (2014a) and include the lagged price to capture the effect of housing appreciation on housing demand.

We chose the explanatory variables basing on the previously discussed empirical literature and the detailed statistical and graphical analysis of house prices and their drivers in the Polish house market presented in detail in NBP (2013). As a detailed analysis of the above

mentioned house price determinants by Baldowska et al. (2013) shows, a considerable part of these variables follows an upward trend only, rather than to display fluctuations likely to explain the ups and downs in prices and the boom period in the housing market. Moreover, some time-series show a strong collinearity. If those variables were included, they would lead to spurious regressions. After running numerous regressions, we decided to include the following explanatory variables: lagged price, average wages in the enterprise sector, the unemployment rate, the loan availability, weighted interest rate and construction costs<sup>4</sup>. All monetary variables are deflated with the CPI in order to exclude the inflation trend. For each market, we use local explanatory variables.

In all the regressions we use the logarithms of the above-mentioned variables, which helps us to capture the non-linear relationships between price changes and the explanatory variables. As a theoretical model of house prices in Augustyniak et al. (2014a) shows, there are important nonlinear relationships between house prices and fundamental variables that need to be accounted for empirically. We apply the fixed effects regression method<sup>5</sup> with robust standard errors (bootstrapped 1 000 times).

---

<sup>4</sup> The loan availability is calculated under the assumption that loans denominated in zloty and in foreign currency were granted during the period 2005-2011, whereas in the remaining years only zloty denominated loans were granted.

<sup>5</sup> The choice of the fixed effects regression model has theoretical foundations. This method is used when the selected sample is not a random, but represents the entire population. Moreover, the economic analysis of individual markets, presented by Baldowska et al. (2014), shows that each market has a unique character, which practically does not change over time. The fixed effects method makes it possible to exclude this fixed element which is impossible to detect with any of the included explanatory variables, and would be erroneously attributed to the error term of the model. We also run the Hausman test. It showed that the random effects model can be used, however, the results of this test can be considered reliable only if a much bigger number of observations (at least 20-30 observations in a time-series) is present.

We are aware of the problems that the fixed effects regression can cause in dynamic panels (see Kiviet, 1995 and Flannery and Hankins, 2013). We therefore also applied the *bias corrected least squares dynamic panel data estimator* proposed by Bruno (2005). We found that the regression results are nearly identical to the FE ones, thus we stayed with the FE regression. We also tested the Bruno estimator, including time dummies. In this case we face the problem that interest rates or construction costs, which are identical for all cities, drop out of the regression. It is crucial to capture the effect of the interest rate and the construction costs separately, therefore we had to skip the time dummies.

Home prices and most explanatory variables are non-stationary, yet, the Pesaran test (2004) shows that the error terms are not correlated and we find that the error terms are stationary. This allows us to conclude that the regression models are correctly specified. The tests are repeated for every specification that is presented in this paper.

The empirical analysis is divided into three units. The first and second deal with the primary market and secondary (existing) market analysis, respectively. In the third part we explain prices in the primary market using data from the secondary market and vice versa. We also investigate the price spillover effects from the Warsaw market to other local markets.

## **2.1 Analysis of the primary market**

In the first regression, we explain the transaction price per square meter of housing with its lagged price, the rate of unemployment, building costs and the average wage in the enterprise sector. All explanatory variables are significant at the 1 percent level and the regression seems to be

well specified. We find that the current price depends on the past price with an elasticity of around 40 percent. This finding confirms the hypothesis that the appreciation of housing adds to housing demand. A decline in the unemployment rate, thus growth in employment leads to higher housing demand that translates into a rising house price. The elasticity of price changes in response to changes in the unemployment rate is of around -30 percent and the parameter is highly significant. Increases in real wages have a strong positive effect and their elasticity is around 70 percent. Moreover, we included the growth rate of construction costs. Because construction costs depend on the lagged construction level and its price there might be an endogeneity problem. We therefore use the first differences of the construction costs instead of its levels and we find that developers ask for higher prices if construction costs increase.

In regression (2) we add to the core model (1) the weighted interest rate<sup>6</sup> for housing loans, which determines the loan cost. All the core parameters remain the same, but the interest rate does not have any effect on the price level. In order to analyse the potential effect of loans we include the available loan in regression (3). This variable is directly constructed from the income and the weighted interest rate, thus we need to drop those variables in this regression. Contrary to our expectations, the loan availability has no effect on house prices in the primary market. During the bust period interest rates remained low and wages continued to increase, but banks slowed down loan disbursements and house prices started to revert to pre-boom levels slowly. We could expect that the loan availability has the same positive impact as wage increases have, but it turns out to be insignificant.

---

<sup>6</sup> The interest rate is the weighted sum of the PLN and FX loan interest rate, weighted by the quarterly volume of loans in domestic and foreign currency.

Housing in the primary market is bought at least to some extent with cash (either from savings or from selling another dwelling, see NBP 2013), therefore the loan availability seems to be less important. Interestingly, in this specification the past price has an even higher effect on the current price. One possible explanation is that banks were giving out loans basing on positive expectations as concerns house price movements. The second explanation is that when the interest rate increased (especially in the boom period) people took their loans for longer and longer periods, i.e. for 25 or 30 years instead of 20 years. This extended repayment period allowed households to keep quite low monthly instalments even though interest rates were rising. Another explanation is the fact that developers adjusted prices downwards much slower than they increased them. Therefore, rising interest rates did not lead to strong price declines and therefore they turn out to be insignificant. In all other modifications of core model (1) the interest rate turned out as unimportant.

We also present modifications of the second model in which we use the growth rates of some variables instead of their levels. When we substitute the unemployment rate with its growth rate (regression 4), the results do not change much. However, we observe that the lagged price has a higher coefficient 0,66 and also reacts stronger to growing construction costs (coefficient of 0,85). One conclusion is that the unemployment level has a stronger effect on house prices than its dynamics have. Finally, we substitute the wage levels with their growth rate (regression 5) and rerun the regression. While the lagged price, unemployment level and credit costs have the same effect as in regression (2), the growth rate of wages has basically no effect on prices. Most likely the fluctuations of wages have no

effect on the market as potential buyers focus on the general wage movements and base their expectations on them.

Table 1. Regression results for the primary housing market in 17 cities, 2002-2013

	(1)	(2)	(3)	(4)	(5)
L_price_prim					
L_price_prim	,3819***	,4001***	,5488***	,6689***	,5422***
L1	[,1017]	[,1044]	[,0583]	[,0488]	[,0802]
L_unemployment	-,3350***	-,3168***	-,3409***		-,3592***
	[,0693]	[,0711]	[,0574]		[,0698]
L_unemployment				-,2084***	
D1				[,0480]	
L_constr_costs	,4266***	,4671***	,2777*	,8525***	,3071
D1	[,1691]	[,1780]	[,1692]	[,1435]	[,2109]
L_wages	,7023***	,6997***		,7499***	
	[,2026]	[,1703]		[,1943]	
L_wages					,0107
D1					[,4232]
L_credit_costs		-,0315		-,0739	-,0346
		[,0766]		[,0651]	[,0699]
L_loan_availability			,1525		
			[,0948]		
constant	0,1815	,0708	2,4835*	-3,0822**	-,0346
	[1,1286]	[1,0060]	[1,3853]	[1,2656]	[,0699]
No. Obs.	187	187	187	187	187
R-sq. within	0,8735	0,8739	0,8643	0,8628	0,8624
between	0,6240	0,6373	0,7675	0,8242	0,7920
overall	0,7503	0,7585	0,8176	0,8445	0,8344

Level of significance: 10% - \*, 5% - \*\*, 1% - \*\*\*; robust standard errors in brackets.

The analysis of the price determinants in the primary market leads us to conclude that house prices depend on the past price, thus the housing appreciation has an impact on housing demand. Moreover, the

unemployment level has a negative effect on house prices, while they rise with the wage level. These findings seem to confirm that fundamental variables have a major impact on house prices and that buyers form expectations on past prices. The loan availability or the cost of credit measured with the weighted interest rate does not have any impact on prices. While this is quite counterintuitive, it might result from the fact that a significant share of primary market housing is bought with cash, thus loans might be relatively less important. Finally we observe that rising construction costs make developers increase prices and to transfer the higher costs to the buyers.

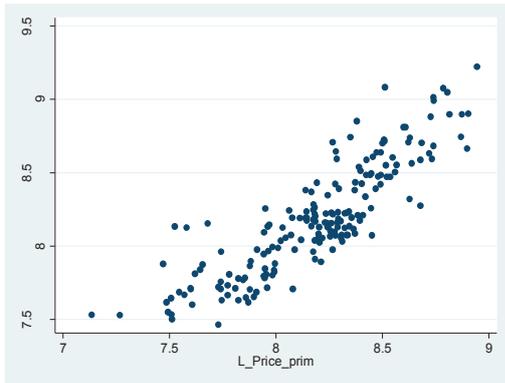
We also perform an analysis of the actual observed prices and predicted prices in individual cities, which confirms that the model is well specified (graphs 3-6). To improve the visibility of the results, we divide the cities into eight large and nine smaller ones<sup>7</sup>. The fit is quite good in most cities and only in the case of Katowice we observe a significant shift between actual and predicted data (see graph 6). Katowice is a quite small market in the whole Silesian agglomeration, shows a very high average wage level and a low unemployment rate. The city has a large housing stock in relation to its inhabitants. Many people who work in Katowice live outside of the city. The housing demand in the city is well satisfied and therefore the real housing demand is lower than the predicted one. In consequence the house price in Katowice is overestimated, if we base it on macroeconomic

---

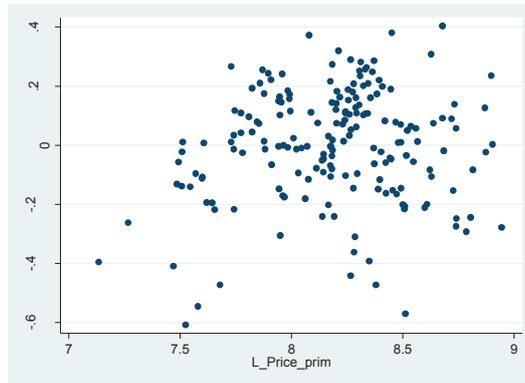
<sup>7</sup> The division was based on the number of inhabitants in a particular city. "Large" cities, i.e. with population exceeding 400 thousand inhabitants include: Gdańsk and Gdynia (as one big market), Kraków, Łódź, Poznań, Szczecin, Warsaw and Wrocław. The group of "small" cities, i.e. with population of less than 400 thousand inhabitants include: Białystok, Bydgoszcz, Katowice, Kielce, Lublin, Olsztyn, Opole, Rzeszów and Zielona Góra.

fundamentals. We also tried to run regressions 1-5 without Katowice but the results do not alter significantly in the remaining cities.

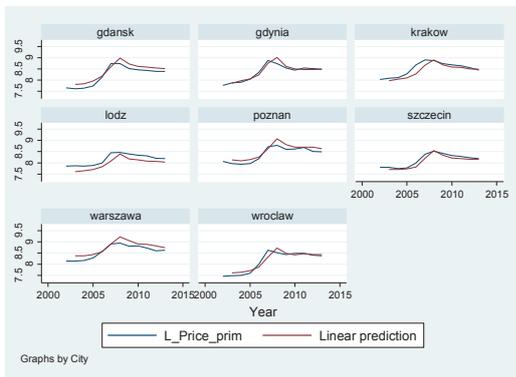
Graph 3 Predicted and actual observations for Model (1) – Primary market



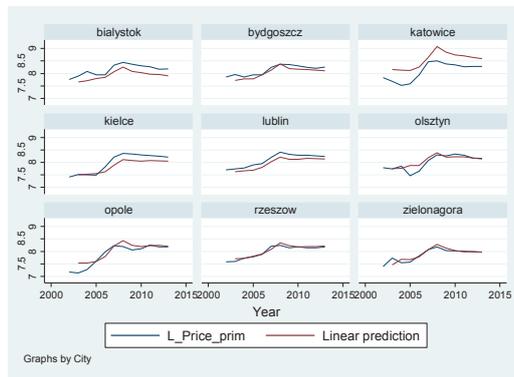
Graph 4 Error and real data for Model (1) – Primary Market



Graph 5 Predicted and actual observations for Model (1) for „big cities”



Graph 6 Predicted and actual observations for Model (1) for „small cities”



## 2.2 Analysis of the secondary market

The next step in our analysis is the secondary market. The estimation of various model specifications leads us to the same models that we apply to the primary market. The model for the primary market (Model 1) seems to be well suited for the secondary market too, has even better R-sq values and most explanatory variables are significant at the 1 percent level. In model 6

the current price depends positively on the lagged price with an elasticity of around 50 percent. The arising unemployment level decreases price levels with a parameter of around -0,3 while they rise with wages. Surprisingly, the growth of construction costs has a significant impact on house prices in the secondary market, too. One possible explanation is that if construction costs in the primary market rise, consumers tend to substitute primary market housing with cheaper housing from the secondary market. This pushes secondary market prices upwards.

We extend the model and include the cost of credit in regression 7. While the remaining parameters do not change significantly, the cost of credit has a significant negative effect on house prices. This finding is in line with the fact that a large share of secondary market housing is financed with a loan. Rising interest rates make loans more costly and less available and therefore curb the price growth in the secondary market.

In all specifications the current price depends a little bit stronger on the past period price than it happens in the primary market. This finding can be explained in the following way. On the primary market supply is quite inelastic, thus increasing demand leads to strong price increases. On the secondary market, as prices rise, more and more people decide to put their house on the market. This way supply is more elastic than in the primary market and the price increases slower.

Table 2. Regression results for the secondary housing market in 17 cities, 2002-2013

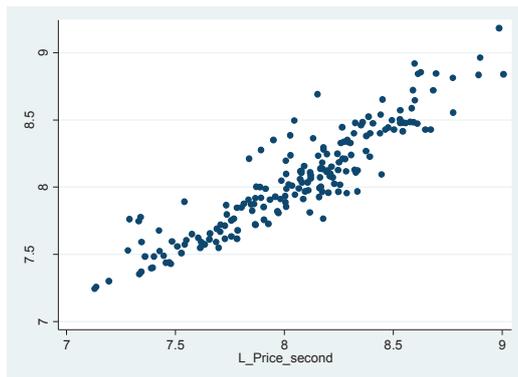
	(6)	(7)	(8)	(9)	(10)
L_price_sec	,5166***	,5794***	,6190***	,7741***	,6504***
L1	[,0451]	[,0469]	[,0309]	[,0295]	[,0369]
L_unemployment	-,3047***	-,2309***	-,2716***		-,2520***
	[,0483]	[,0504]	[,0530]		[,0587]
L_unemployment				-,1909***	
D1				[,0318]	
L_constr_costs	,5587***	,7091***	,4811***	,8776***	,5947***
D1	[,1598]	[,1511]	[,1488]	[,1216]	[,1847]
L_wages	,3642**	,3725***		,3610***	
	[,1459]	[,1232]		[,0865]	
L_wages					,1019
D1					[,3655]
L_credit_costs		-,1461***		-,1542***	-,1405***
		[,0390]		[,0300]	[,0409]
L_loan_avaliability			,2218***		
			[,0507]		
constant	1,6398	1,1719	,8552	-,7481	3,5848***
	[,9936]	[,8312]	[,7819]	[,5535]	[,3701]
Nr. Obs.	187	187	187	187	187
R-sq within	0,8990	0,9067	0,9010	0,9032	0,9042
between	0,7494	0,8044	0,8084	0,9472	0,8950
overall	0,8384	0,8672	0,8612	0,9180	0,9007

Level of significance: 10 % - \*; 5 % - \*\*; 1 % - \*\*\*; robust standard errors in brackets.

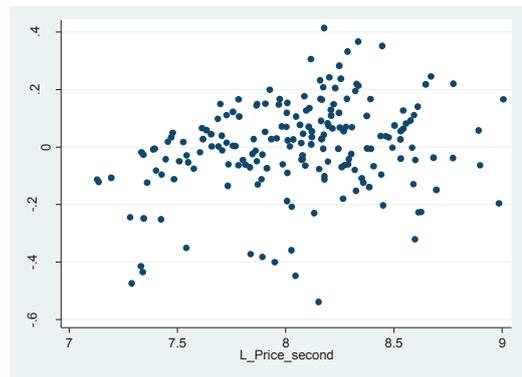
The graphical analysis of the predicted and actual observations as well as the error terms, presented in graphs 7-10, indicates that the model for the secondary market is well specified and suitably explains the price changes in the local markets. Again the results for Katowice are overestimated and the same explanation as for the primary market applies. Moreover, the prices for Bialystok, Lublin and Kielce are systematically underestimated. Those cities have a very high unemployment rate and quite low official wage levels. We assume that there is a significant share of people

who work unofficially or work in larger cities or even abroad and buy housing with a salary that is higher than the city average. Another important factor is the unofficial trade with eastern countries, which boosts the income of people who live in those cities.

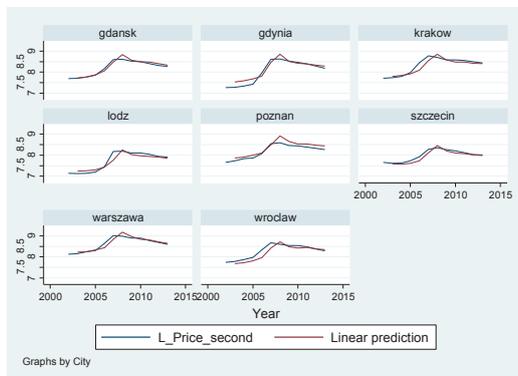
Graph 7 Predicted and actual observations for Model (6) – Secondary market



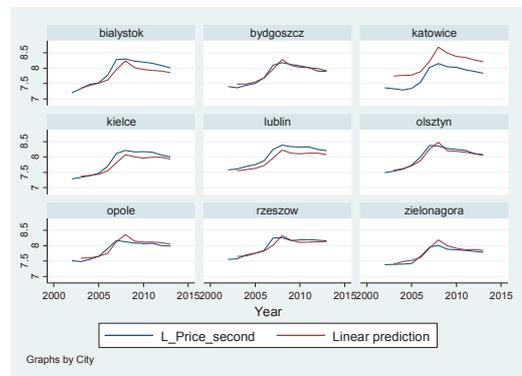
Graph 8 Error term and actual observations for Model (6) – Secondary Market



Graph 9 Predicted and actual observations for Model (6) for „big cities”



Graph 10 Predicted and actual observations for Model (6) for „small cities”



### **2.3 Interdependence between the primary and secondary market and spillover effects**

In the third section of this paper we investigate the mutual interdependence between the primary market prices and secondary market prices (Model 11 and 12). We also analyse the price spillover effects from Warsaw to other local cities for both markets (Model 13-16).

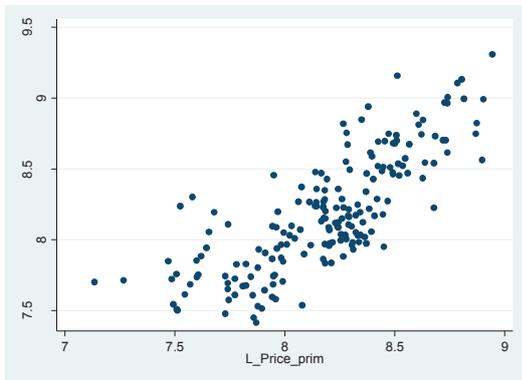
We use the same set of explanatory variables as in regression 1 and 6 to explain prices in the primary or secondary market, but add the lagged price of the other market. The most important empirical result is that lagged prices in the secondary market help to explain the prices in the primary market, but it does not work the other way around. This means that price signals from the secondary market spread to the primary market. The explanation of this result follows the empirical observations of the market described in NBP (2013) and also bases on the theoretical model of interactions between the primary and secondary market proposed by Augustyniak et al. (2014a). Firstly, if demand in the secondary market is excessive, at some point the housing demand can be only satisfied with newly constructed housing from the primary market, thus its prices increase. Secondly, housing developers are able to use marketing techniques to rise prices in the primary market above a value that fundamentals could explain. On the other hand, when prices in the primary market rise, they have little effect on the secondary market. The secondary market is huge and along with rising prices more and more owners decide to sell their housing, which has a stabilizing effect on price increases.

Table 3 Primary Market explained with secondary market and vice versa; Price spillover effect from the Warsaw market to other local markets, 2002-2013

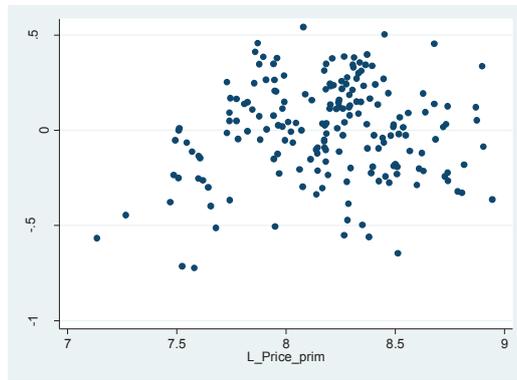
L_price_prim	(11)	--	(13)-W	--	(15)-W	---
L_price_sec	--	(12)	--	(14)-W	---	(16)-W
L_price_prim L1	--	,1218 [,0796]	,3892*** [,0924]	--	--	--
L_price_sec L1	,1905** [,0973]	--	--	,5240*** [,0435]	--	--
L_unemployment	-,4246*** [,0742]	-,4793*** [,0654]	-,3345*** [,0637]	-,3050*** [,0498]	-,2823*** [,0830]	-,1291* [,0702]
L_constr_costs D1	,1618 [,1609]	,1719 [,2206]	,4270*** [,1630]	,5632*** [,1607]	-,1933** [,0927]	-,1499 [,1491]
L_wage	1,0109*** [,2394]	1,3322*** [,2117]	,6875*** [,1774]	,3545** [,1445]	,8634*** [,1632]	,7387*** [,1264]
L_price_prim_WAW					,4713*** [,0882]	
L_price_sec_WAW						,7456*** [,0740]
constant	-,5003 [,15176]	-2,5321 [,15733]	,2472 [,10966]	1,6583* [,09968]	-2,1389 [,17351]	-4,0062*** [,12425]
No. Obs.	187	187	176	176	176	176
R-sq within	0,8492	0,8419	0,8731	0,9013	0,8637	0,9209
between	0,4611	0,3499	0,5028	0,6200	0,2340	0,1067
overall	0,6246	0,5468	0,7264	0,8130	0,6432	0,6734

Level of significance: 10 % - \*, 5 % - \*\*, 1 % - \*\*\*; robust standard errors in brackets.

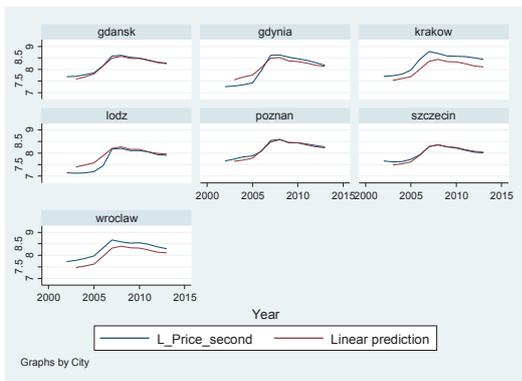
Graph 11 Predicted and actual observations for Model (11)



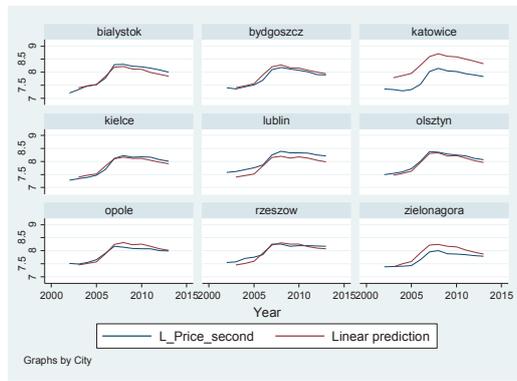
Graph 12 Error term and actual observations for Model (11)



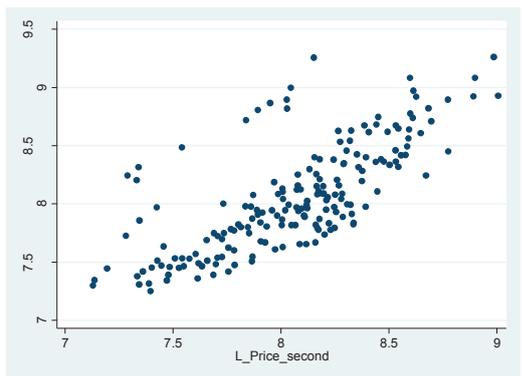
Graph 13 Predicted and actual observations for Model (11) for „big cities”



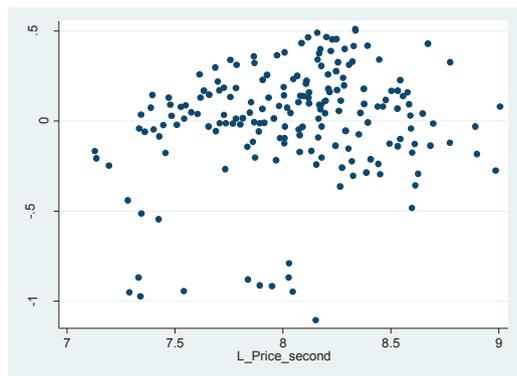
Graph 14 Predicted and actual observations for Model (11) for „small cities”



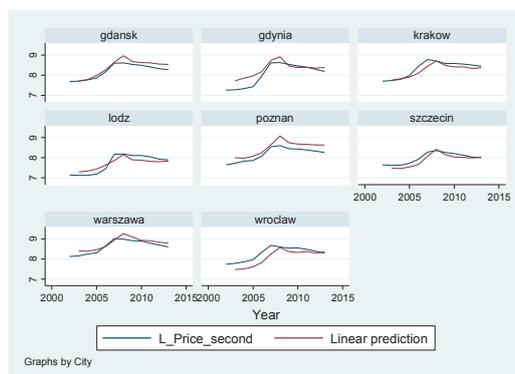
Graph 15 Predicted and actual observations for Model (12)



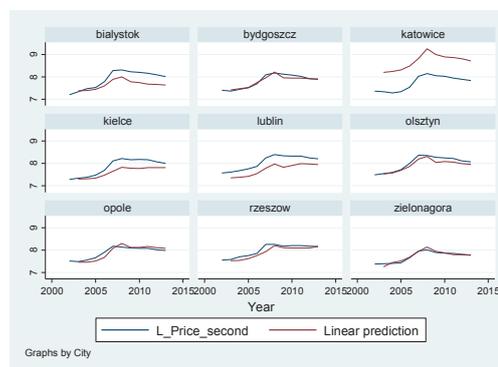
Graph 16 Error term and actual observations for Model (12)



Graph 17 Predicted and actual observations for Model (12) for „big cities”



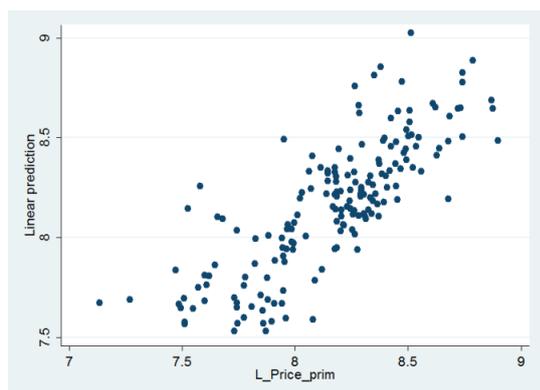
Graph 18 Predicted and actual observations for Model (12) for „small cities”



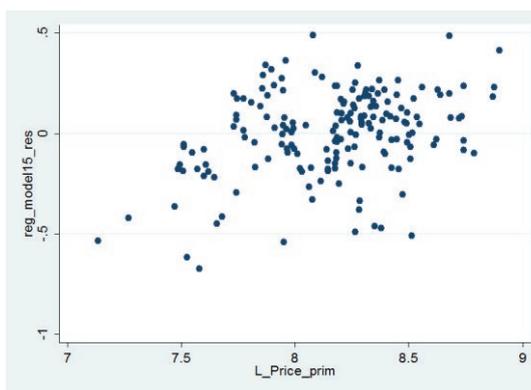
In the last set of regressions, we investigate whether there is a contagion effect from rising prices in Warsaw that spread to other local markets. In regressions 13-16 we include all cities but Warsaw. First, we run regressions 13 and 14 that replicate regressions 1 and 6, respectively, and find that the results stay very similar. Basing on this specification, we add the current price of housing in the Warsaw market as an explanatory variable. We have to exclude the lagged price in the local market, because the price levels are collinear. If we include the Warsaw price and the lagged price we would get a spurious regression. We include the price in Warsaw in the same year, because we expect that news about rising prices spread from Warsaw to other cities in the same period. Indeed the regression results of regression 15 and 16 indicate that the Warsaw market has a significant impact on the primary and secondary market. In the primary market the effect of the unemployment level remains unchanged in comparison to regression 13, while the wages have a slightly higher effect on prices. The rise of prices in Warsaw affects the local price with a parameter of 0,47. Surprisingly, the effect of the increase of construction costs has a negative sign. It is possible that the effect of the Warsaw price and the wages is overestimated and therefore the construction costs obtain a negative sign. In the secondary market we find that impact of the Warsaw market price has a parameter of 0,74. The effect of the unemployment level is lower than in regression 14, but still remains significant. The elasticity of prices in relation to wages is around two times

higher than in regression 14. As in the case of the primary market the effect of rising construction costs is negative, although insignificant. The regression results, measured with the distribution of the error terms seem to be better for the primary market than for the secondary market.

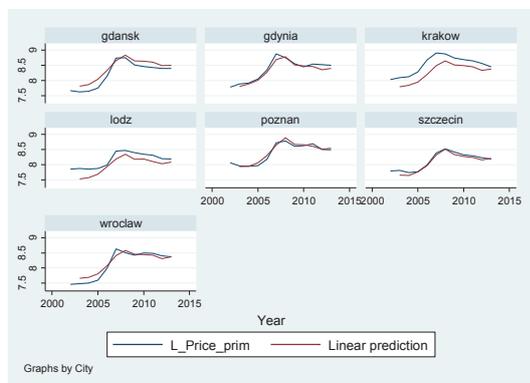
Graph 19 Predicted and actual observations for Model (15)



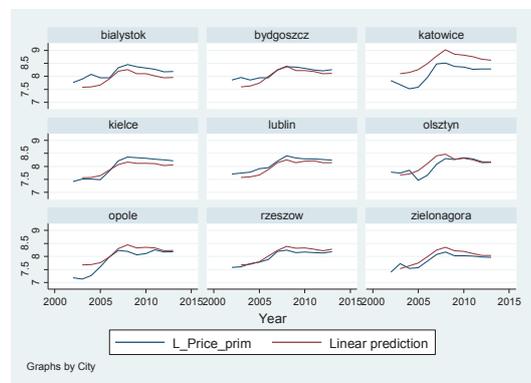
Graph 20 Error term and actual observations for Model (15)



Graph 21 Predicted and actual observations for Model (15) for „big cities”



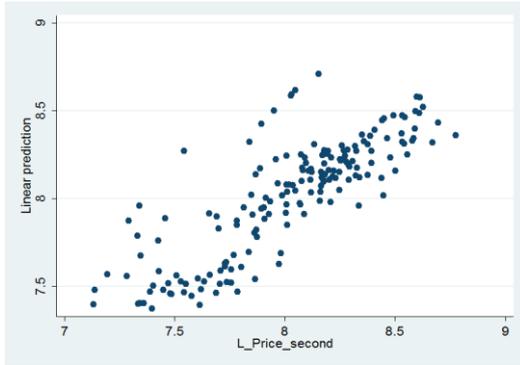
Graph 22 Predicted and actual observations for Model (15) for „small cities”



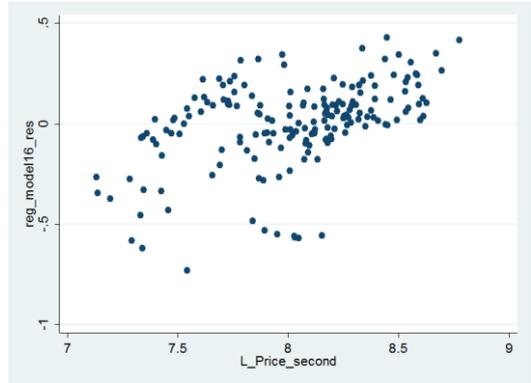
When we include the spillover effect of the Warsaw prices to local markets, the regression results for Białystok, Kielce and Lublin improve significantly. While in regression 11 and 12 we see that the prices in those markets, both in the primary and the secondary market, are underestimated, now the predicted prices is very close to the empirical one. This confirms our previously mentioned assumption that wages in those three cities are underreported. The price development in Warsaw is

a proxy for performance of the overall economy, therefore it explains the price developments in those small cities well.

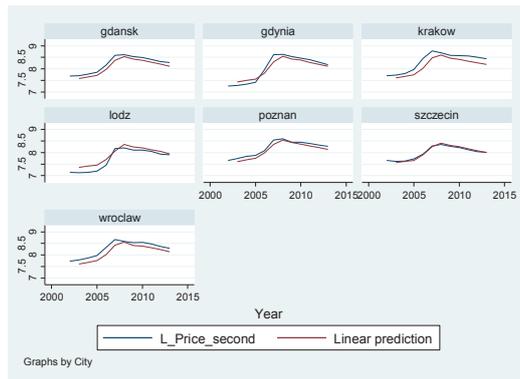
Graph 23 Predicted and actual observations for Model (16)



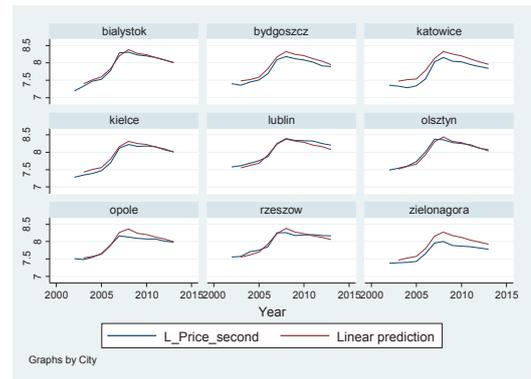
Graph 24 Error term and actual observations for Model (16)



Graph 25 Predicted and actual observations for Model (16) for „big cities”



Graph 26 Predicted and actual observations for Model (16) for „small cities”



### 3. Conclusions

The analysis confirms that transaction prices in the primary and secondary market in 17 largest cities in Poland depend on fundamental variables such as wages, the rate of unemployment and construction costs. Moreover, we confirm empirically that the appreciation of housing has a positive effect on the price level in the next period. We conclude that appreciation affects the utility function of the household and impacts on its optimal decision, which strengthens the theoretical results presented in Augustyniak et al. (2014a).

We also find that that prices in the secondary market help to explain the prices in the primary market, however this relationship cannot be inverted. This finding confirms that the demand shock goes first to the secondary market and then to the primary market. Potential buyers try to find housing in the secondary market first, because it is usually cheaper. But if demand exceeds supply, they have to go to the primary market, which increases the prices in the primary market. However, price rises in the primary market have little effect on the prices in the secondary market. Most likely, if prices in the secondary market start to increase, more and more owners put their housing on sale, thus supply increases and becomes more flexible. This prevents price bubbles in the secondary market.

Finally, we confirm the existence of a contagion effect of price rises in the Warsaw market, that spill over to the other largest cities. As news of rising prices in the capital city spread over the country, developers and house owners in local markets adjust their price expectations. We find that the spillover effect is stronger in the secondary market than in the primary one. Intuitively, one would expect that the primary markets, which are often operated by the same developers behave similarly. However, it turns out

that the price of the secondary market in Warsaw is a good proxy for the economic performance of the whole economy. It helps to properly estimate the house prices, which in some cases, most likely due to underreported wages and unofficial employment, were underestimated.

Our results should be useful for other countries, both highly developed and for those who are still emerging. To our best knowledge this is the first study that investigates the quite complicated relationships between the primary and secondary market for an Eastern European Country, which allows to determine how a demand shock affects different segments of the market and shows which fundamental variables have a significant impact on the primary or secondary market. The results can be especially useful for emerging markets, which are faced with a high demand for housing and where the existing housing stock is quite depreciated. The results also indicate that the primary and secondary market, which have different price and demand determinants, need to be monitored carefully and analysed separately.

**Literature:**

- André, C. (2010), *"A Bird's Eye View of OECD Housing Markets"*, OECD Economics Department Working Papers, No. 746, OECD Publishing.
- Andre, C., Gil-Alana, L. A., and Gupta, R. (2013). *"Comovement in Euro Area Housing Prices: A Fractional Cointegration Approach"*. University of Pretoria Working Paper: 2013-59.
- Andrews, D. (2010), *"Real House Prices in OECD Countries: The Role of Demand Shocks and Structural and Policy Factors"*, OECD Economics Department Working Papers, No. 831, OECD Publishing.
- Augustyniak, H., K. Gajewski, J. Łaszek and G. Żochowski (2012), *"Real estate development enterprises in the Polish market and issues related to its analysis"*, MPRA Paper 43347.
- Augustyniak, H., J. Łaszek, K. Olszewski and J. Waszczuk (2014a), *"Modelling of cycles in the residential real estate market – interactions between the primary and the secondary market and multiplier effects"*, *Kwartalnik Nauk o Przedsiębiorstwie* 2/2014 (31).
- Augustyniak, H., J. Łaszek, K. Olszewski and J. Waszczuk (2014b), *"To rent or to buy – analysis of housing tenure choice determined by housing policy"*, *Ekonomia* V33, 31-54.
- Augustyniak, H., J. Łaszek, K. Olszewski and J. Waszczuk (2014c), *"Housing market cycles – a disequilibrium model and its application to the primary housing market in Warsaw"* *Ekonomia*, forthcoming.
- Baldowska G., B. Myszkowska and R. Leszczyński (2014). *"Convergence and differentiation processes in local markets and structural changes (comparison of 16 markets in Poland)"* NBP Working Paper No. 174.

- Bruno, G. S. F. (2005). "Estimation and inference in dynamic unbalanced panel-data models with a small number of individuals," *Stata Journal*, StataCorp LP, vol. 5(4), 473-500.
- Burnham, J. B. (1972), *Private Financial Institutions and the Residential Mortgage Cycle, with Particular Reference to the Savings and Loan Industry*, In: Board of Governors of the Federal Reserve System, *Ways To Moderate Fluctuations in Housing Construction*.
- Cameron G., J. Muellbauer and A. Murphy (2006) "Was There A British House Price Bubble? Evidence from a Regional Panel," *Economics Series Working Papers 276*, University of Oxford, Department of Economics.
- Chang, C.-O. and C.W.R. Ward (1993), "Forward pricing and the housing market: the pre – sales housing system in Taiwan", *Journal of Property Research* V 10, 217-227.
- De Bandt, O., K. Barhoumi and C. Bruneau (2010), *The international transmission of house price shocks*, In: *Housing Markets in Europe*, Springer Berlin Heidelberg, 129-158.
- DeFusco, A., Ding, W., Ferreira, F., and Gyourko, J. (2013). *The Role of Contagion in the Last American Housing Cycle*. Mimeo.
- Dittmann, I. (2013). "Primary and Secondary Residential Real Estate Markets in Poland—Analogies in Offer and Transaction Price Development", *Real Estate Management and Valuation* 21.1: 39-48.
- Ferreira, F., and Gyourko, J. (2011). "Anatomy of the beginning of the housing boom: US neighborhoods and metropolitan areas, 1993-2009", *NBER Working Paper* 17374.
- Flannery, M. J., and Hankins, K. W. (2013). "Estimating dynamic panel models in corporate finance". *Journal of Corporate Finance*, 19, 1-19.

- Fry, J. M., 2009. "Bubbles and contagion in English house prices", MPRA Paper 17687.
- Igan, D. and P. Loungani (2012), "Global housing cycles", IMF Working Paper No. 12/217.
- Jud, G.D. and D.T. Winkler, 2002. "The Dynamics of Metropolitan Housing Prices," Journal of Real Estate Research, American Real Estate Society, vol. 23(1/2), 29-46.
- Kiviet, J.F., (1995). "On bias, inconsistency, and efficiency of various estimators in dynamic panel data models". Journal of Econometrics 68, 53–78.
- Łaszek J. (2013) "Housing and consumer theory", MPRA Paper 52599.
- NBP (2013) "Report on the situation in the Polish residential and commercial real estate market in 2012", National Bank of Poland
- Pesaran, M. (2004), "General Diagnostic Tests for Cross Section Dependence in Panels", Cambridge Working Papers in Economics No. 0435, Faculty of Economics, University of Cambridge.
- Vansteenkiste, I. and P. Hiebert, 2011. "Do house price developments spillover across euro area countries? Evidence from a global VAR", Journal of Housing Economics, 20(4), 299-314.
- Waszczuk J. (2013), *Przestrzenne modele ekonometryczne na przykładzie warszawskiego rynku nieruchomości mieszkaniowych*, Bezpieczny Bank.
- Widłak M. (2013), "Study of factors that differentiate housing prices and the possibility of their use at NBP", in: Report on the situation in the Polish residential and commercial real estate market in 2012, NBP.
- Widłak, M. J. Waszczuk and K. Olszewski (2014). *Spatial and hedonic analysis of house price dynamics in Warsaw*. Mimeo.



---

## 13. Non-listed real estate funds: leverage and macroeconomic effects

Franz Fuerst, Wayne T. Lim, George A. Matysiak



Franz Fuerst – University of Cambridge. Department of Land Economy, 16–21 Silver Street, Cambridge CB3 9EP, United Kingdom, ff274@cam.ac.uk

Wayne T. Lim – University of Cambridge. Queens' College, Silver Street, Cambridge CB3 9ET, United Kingdom, wl306@cam.ac.uk

George A. Matysiak – Master Management Group, Warsaw & Krakow University of Economics, gmatysiak@gmail.com

The authors would like to thank Casper Hesp of INREV and an anonymous referee for their detailed comments.

---

## Article 13 Contents

1. Introduction .....	77
2. Empirical evidence on individual performance drivers .....	77
3. Panel data analysis of fund returns .....	82
4. Conclusions and further work .....	90
Appendix .....	91
References .....	92

### *Abstract*

*This research investigates the key drivers of European non-listed real estate funds performance. In particular, it seeks to understand the extent to which stock-selection, management skills, gearing, fund size, competing asset classes and macroeconomic performance contribute to a fund's total return. Our attribution of fund performance is primarily based on average return data in the markets to which the fund is exposed. Using a proprietary database compiled by INREV covering the period 2001-2012, we analyse the performance characteristics of non-listed funds. Tracking fund performance and its drivers through the years of rapid expansion of the sector, followed by the sharp contraction during the global financial crisis, allows us to establish whether performance drivers have a differential impact during market upturns and downturns. We find that fund performance during the study period was predominantly driven by the underlying market and sector exposure of a fund, gearing, fund size, GDP growth, and competing asset classes.*

Keywords: real estate, private equity, non-listed funds, leverage, macroeconomic effect

## 1 Introduction

Alternative investments are recognized for their portfolio diversification properties and their perceived low correlation with stock markets. In particular, real estate investments have been gaining popularity over several years and make up a substantial portion of institutional portfolios (Georgiev, Gupta, & Kunkel, 2003). Investors looking to gain exposure to underlying property assets may look to property investment vehicles such as non-listed real estate funds (NLREFs).

Investors recognize the benefits of NLREFs and the sector has experienced explosive growth over the last decade (Brounen, Veld, & Raitio, 2007). This growth has helped establish NLREFs as a major investment vehicle (Fuerst & Matysiak, 2013). NLREFs facilitate cross-border property investments, making it the preferred conduit for investors who are looking for real estate exposure outside their home markets (Baum & Farrelly, 2009).

## 2 Empirical evidence on individual performance drivers

While the existing literature on NLREFs is relatively sparse, studies on listed real estate funds or the general fund literature can also provide important clues as to the main drivers of non-listed financial performance. The following review discusses each of these drivers separately.

### *Fund Size*

An oft-debated topic revolves around the relationship between fund size and performance. Fund size and fund performance have been found to be inversely associated in several studies (Chen, Hong, & Kubik, 2004). Yan (2008) expands the research by considering liquidity as a control variable and finds that the negative correlation between fund size and performance is more evident among funds that hold less liquid portfolios. Further studies demonstrate liquidity as an important factor in the inverse fund size performance relationship (Chanet al., 2009 and Yang, S.-F., & Wu, T.-H. 2012). Lowenstein (1997) postulates that a large asset base erodes fund performance because of transaction costs associated with liquidity. Conversely, there are advantages to larger funds as well, most significantly, having more resources for research, shared fixed costs and lower expenses (Fredman & Wiles, 1993). In a similar vein, Latzko (1999) finds that the elasticity of fund expenses with respect to fund assets is significantly less than 1, indicating there are scale economies in fund administration.

### *Gearing*

Independent of fund size, fund managers looking to enhance returns will take on some form of gearing. Leverage enhances returns by reducing the amount of equity investment required for a particular asset. It allows equity investors to magnify the amount of underlying capital and transcend capital constraints to achieve more diversification.

Furthermore, leverage can minimise agency costs. High gearing can motivate managers to make better investment decisions for fear of losing control or reputation given the smaller financial buffer for error (Grossman & Hart, 1982). Leverage can also discipline managers since the debt service burden reduces free cash flow and limits managers' scope to engage in personal pursuits (Jensen, 1968).

On average, real estate debt appears to be significantly higher than is the case for other asset classes. The main reason for higher leverage is that the loan is backed by real estate assets as collateral, hence banks are comfortable offering more debt financing.

The 2007 Global Financial Crisis drew further research attention to the role of gearing in the underperformance of funds. Furthermore, empirical evidence suggested that gearing increased the probability of fund failure, especially in times of financial crises (Amoako & Samarbakhsh, 2012). Despite the instrumental role of leverage, there has been limited research on the role of gearing in NLREF performance.

### Data characteristics

The dataset provided by INREV tracks fund level performance of properties of European NLREFs between 2001 and 2012. The dataset includes funds that adopted so-called Core and Value-Added investment styles. The differences in investing styles are summarised in Exhibit 1.

**Exhibit 1: Comparison of Core and Value-added Funds**

	Core	Value-Added
<b>Return Objectives</b>	<ul style="list-style-type: none"> <li>▪ Market-level returns</li> <li>▪ Stable rental income</li> <li>▪ Low-moderate risk</li> <li>▪ Returns mainly from rental income</li> </ul>	<ul style="list-style-type: none"> <li>▪ Expected moderate market outperformance</li> <li>▪ Returns mainly from capital appreciation</li> </ul>
<b>Property Types</b>	<ul style="list-style-type: none"> <li>▪ Office</li> <li>▪ Retail</li> <li>▪ Industrial</li> <li>▪ Residential</li> </ul>	<ul style="list-style-type: none"> <li>▪ Office</li> <li>▪ Retail</li> <li>▪ Industrial</li> <li>▪ Residential</li> <li>▪ May include other property types such as hotel, storage, etc.</li> </ul>
<b>Leverage</b>	<ul style="list-style-type: none"> <li>▪ Upper limit of 50% LTV</li> </ul>	<ul style="list-style-type: none"> <li>▪ Upper limit of 70% LTV</li> </ul>
<b>Other Investment Characteristics</b>	<ul style="list-style-type: none"> <li>▪ Institutional-grade properties, prime locations, high-quality design and tenants</li> </ul>	<ul style="list-style-type: none"> <li>▪ May involve efforts to increase value by refurbishment and/or repositioning</li> </ul>

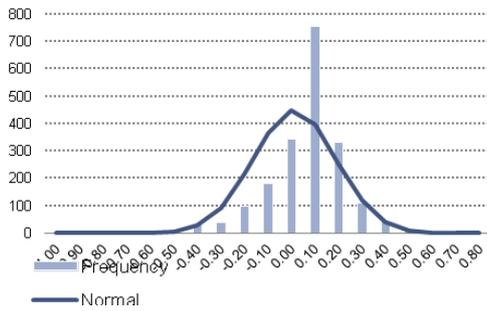
Source: INREV, Fuerst & Marcato, 2009

Fund return, as defined by INREV, is the return investors receive net of management fees. Annual fund returns in the dataset ranged from -80.5% to 62.3%, with a median of 3.7%. The returns are negatively skewed and exhibit leptokurtosis (Exhibit 2). As in the literature review, the fat tails observed here are likely due to the numerous extreme negative fund returns recorded during the Global Financial Crisis.

In order to preserve the sample size and lessen the influence of outliers, Winsorization was applied to fund returns at the 1% level on both tails, replacing extreme values with the closest non-extreme observation (Hastings, 1947). Winsorization maintains the median while shifting the mean closer to the median.

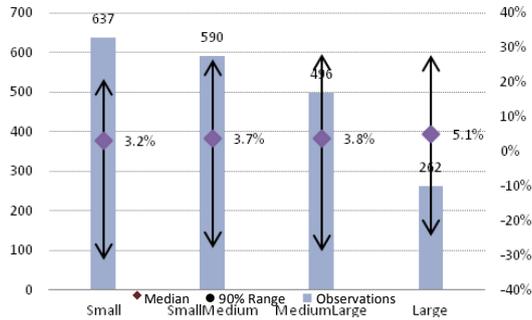
To understand how fund returns vary, they were segmented by Gross Asset Value (GAV) into four categories: “Small” (<€250m), “Small Medium” (>€250m, <€500m), “Medium Large” (>€500m, <€1000m), and “Large” (>€1000m). Fund returns increased with fund size – on average, Small funds returned 3.2% while Large funds returned 5.1% (Exhibit 3). This raises the question if better performing funds attract more capital or if larger funds indeed perform better.

**Exhibit 2: Distribution of Fund Returns, 2001-2012**



Kurtosis 4.53 | Skew: -1.027  
Source: INREV, Authors' calculations

**Exhibit 3: Fund Returns by GAV, 2001-2012**

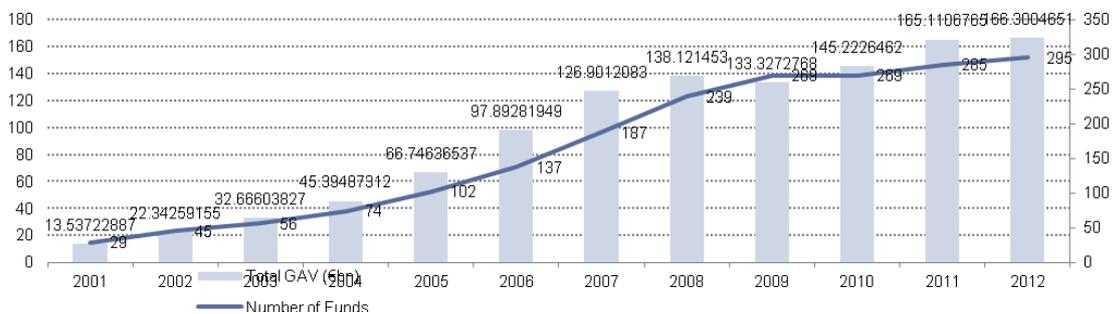


Source: INREV, Authors' calculations

### Gross Asset Value (GAV)

Unlisted real estate funds have grown significantly over the period 2001 and 2012, both in terms of cumulative GAV and number of funds (Exhibit 4). The Global Financial Crisis had reduced the average GAV per fund, and in 2012 this had yet to recover to pre-crisis highs. Diversified funds had a median GAV of EUR 413m, greater than the specialised funds' EUR 344m. One reason is that diversification benefits in real estate can only be realised with larger assets under management because of the large lot sizes and capital intensity of property investments. The Median fund size of core and value-added funds were similar; however, there were numerous large core funds. While median fund size of low gearing and high gearing funds were similar, larger funds had a propensity towards lower gearing. There was also no significant difference in the fund sizes of open-ended and closed-ended funds.

**Exhibit 4: Evolution of Total GAV and Number of Funds, 2001-2012**



Source: INREV, Authors' calculations

### Gearing

Gearing, as defined by INREV, measures fund level gearing. The gearing of highly geared funds remained relatively constant at around 60%, while gearing of the low geared funds had increased in the run up to the Global Financial Crisis (Exhibit 5). The caveat is if underlying asset values had fallen, the equity portion of the capital structure fell and consequently, the level of gearing may appear to have increased. An analysis of subgroups of funds reveals that specialized funds had a lower median gearing of 31.2% compared to the median of 45.7% of diversified strategy funds. Further decomposition reveals that single sector, multi

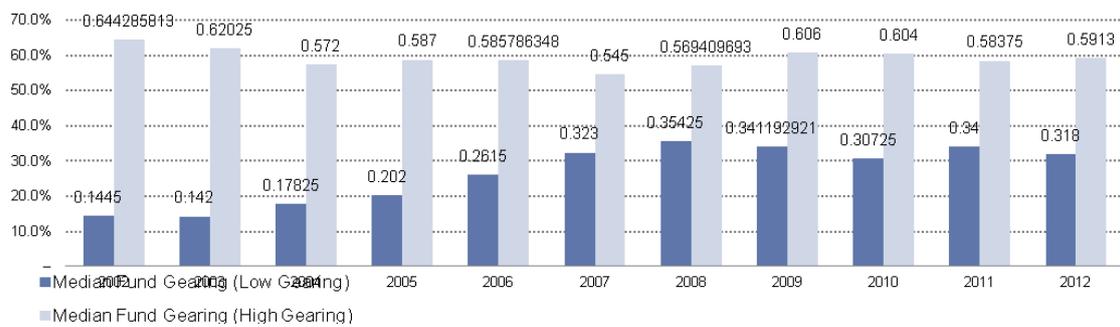
country funds had the highest median gearing of 50.7%. Interestingly, single sector, single country funds had the lowest gearing, but exhibited the best overall performance.

As expected, the riskier value-added funds exhibited higher median gearing of 53%. By structure, closed-ended funds had almost twice the gearing of open-ended funds. Perhaps, the relative certainty of closed-ended structure gives fund managers confidence to take-on more debt, or that banks are relatively more willing to lend. The Median gearing values of Small, Small Medium and Medium Large funds were broadly similar. Large funds, however, had a lower median gearing of 38%. The INREV dataset tracks annual data and provides the spot gearing levels as of year-end. In theory, the net effect of gearing on fund performance for a given year is the result of time weighted gearing levels. Taking into account of gearing levels in the previous period, we calculate gearing as the average gearing over two consecutive years:

$$G_{Avg,t} = \frac{G_{t-1} + G_t}{2}$$

For example, if the gearing for a fund is 20% in December 2006 and 40% in December 2007, the gearing effect on returns for the year of 2007 is calculated as the average between the two data points, 30%. As such, in the empirical analysis that follows, the gearing discussed is the average level of gearing as defined.

Exhibit 5: Median Average Fund Gearing 2002 – 2012



Source: INREV, Authors' calculations

### Weighted Market Return (WMR)

To benchmark fund returns against their underlying market performance, a Weighted Market Return (WMR) was constructed by aggregating the value-weighted returns of individual sector-geographical specific allocations of each fund in each year.

As discussed above, outperformance can be attributed on two levels, namely portfolio structure outperformance and stock selection outperformance. Hence, the WMR would serve as a benchmark, which, if funds beat it, reveals potential managerial stock (property) selection outperformance (Baum A., 2009)<sup>1</sup>.

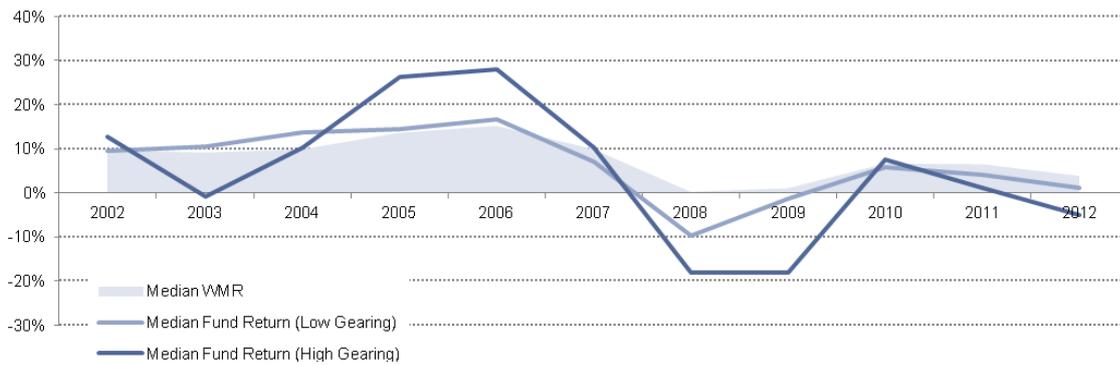
A caveat regarding the comparability of INREV fund returns and WMRs is the potentially different treatment of fees. The commercial property market performance measurer, the Investment Property Databank (IPD), measures performance returns only taking into account asset-level fees, such as property-management costs. In contrast, INREV fund returns are net of all management costs, irrecoverable costs and non-property specific fund management fees. Consequently, the comparability of the two sets of

<sup>1</sup> A full *attribution analysis* would need to be undertaken in order to establish the relative contribution of the asset and stock selection components.

figures may be limited, since the INREV fund returns, on average, would almost certainly be lower given the additional layer of costs associated with fund management fees. Nonetheless, the IPD based WMRs still serve as an important fund return benchmark, offering valuable insights into fund return dynamics.

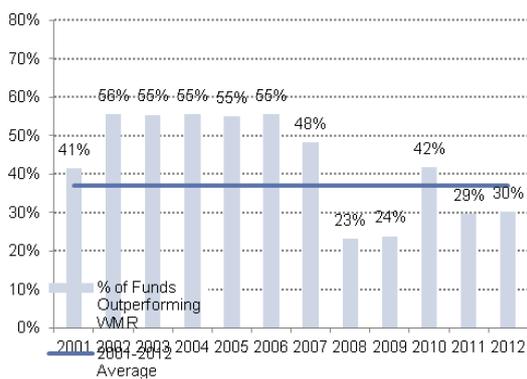
Exhibit 6 shows that mean and median fund returns outperformed their underlying markets over the period 2003-2007. However, fund returns during the Global Financial Crisis were significantly more negative than the WMR. As expected, highly geared funds exhibited greater volatility. However, it cannot be ruled out that the underlying IPD indices used in the construction of WMR suffer from so-called ‘valuation smoothing’ (Brown & Matysiak, 2000), which may overstate/understate the values of commercial properties and of the underlying market performance. The percentage of funds outperforming the WMR and Pan-European IPD index respectively were broadly similar, with more than 50% of the funds outperforming both benchmarks prior to the Global Financial Crisis (Exhibits 7 and 8). However, since the Global Financial Crisis, funds' returns appear to have lagged the benchmarks. The 2001-2012 average percentage of funds outperforming the WMR was 37%, compared to 41% of funds outperforming the Pan-European IPD. This is because WMR outperformance only captures stock level outperformance, while Pan-European IPD outperformance captures both stock level and portfolio structure outperformance.

**Exhibit 6: Median Fund Returns and WMR**



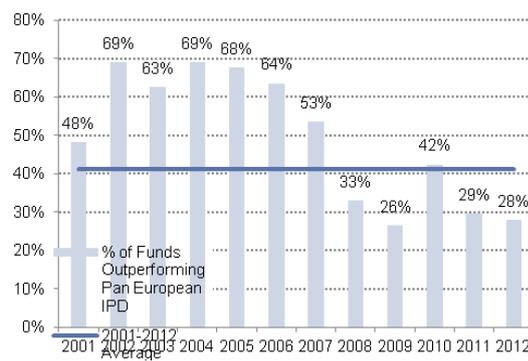
Source: INREV, Authors' calculations

**Exhibit 7: WMR Outperformance by Year**



Source: INREV, Authors' calculations

**Exhibit 8: Pan-European IPD Outperformance by Year**



Source: INREV, Authors' calculations

### 3 Panel data analysis of fund returns

Panel data analysis allows for simultaneous regression analysis of both time-series and cross-sectional research questions. This will enable us to identify and track the drivers of individual fund performance over time. More specifically, panel analysis can be used to capture the dynamics of fund performance in relation to the overall market. There are various types and functional forms of panel data models, including *fixed effects* and *random effects*. A primary advantage in employing fixed effects or random effects models for panel data is the ability of these models to control for omitted variables. Given the more than likely presence/impact of omitted variables this is a major advantage of such models. Fixed effects regression is the model to use when one wants to control for omitted variables whose impact will *differ between cases*, for example, omitted variables having a different impact on investment style returns. If we have reason to believe that some omitted variables *may have the same constant* impact but vary randomly between cases, such as investment styles, we would then model random effects.

We consider both fixed and random effects models for the panel data analysis. Generally, as noted above, fixed effect models are used to control for omitted variables between cases, such as the omitted variables having a *differential* impact on the different investment style returns. However, if the omitted variables are assumed to be uncorrelated with all of the observed variables, and they have an effect which *randomly* varies around a constant between the different investment styles, the random effects model would be suitable.

The *Hausman* test helps ascertain the correct type of model specification in panel data analysis. The fixed-effects model, Equation 1, is based on the assumption that all  $\alpha$  remain constant over time and  $\lambda_t$  coefficients are identical across all funds. As a result, the constant term absorbs the unit effects in the model in the following manner:

$$E(\alpha_i) = E(\lambda_t) = E(\mu_{it}) = 0$$

$$E(\alpha_i \chi_{it}) = E(\lambda_t \chi_{it}) = E(\mu_{it} \chi_{it}) = 0$$

$$Var(\alpha_i) = \sigma_\alpha^2; Var(\lambda_t) = \sigma_\lambda^2; Var(\mu_{it}) = \sigma_\mu^2$$

A two-way error component model is constructed to investigate fund specific effects and cross section effects, both encompassed in the disturbance term. The structure of the fixed-effects model is therefore written as:

**Equation 1**

$$y_{it} = \alpha + \lambda_{it} \beta + \mu_{it}$$

*for*  $i = 1, 2, 3 \dots N; t = 1, 2, 3 \dots, T$

*where*  $\mu_{it} = \lambda_t + v_{it}$

In the regression equation,  $\chi_{it}$  denotes the independent variables that explain the fund return,  $y_{it}$ , for fund  $i$  in year  $t$ . This two-way error component model makes the assumption of unobservable time effects and cross section (individual fund) effects. Here, fund-invariant time effects capture period specific effects that are omitted (Baltagi, 2008).

The annual returns performance analysis considered the following set of variables:

- A. WMR: The weighted market return measures underlying market performance based on fund portfolio selection of geographical and sector for a specific year in percentage.
- B. Gearing: The average of gearing at time  $t$  and gearing at  $t-1$  in percentage.
- C. Gearing Level Dummies:

- Low Gearing Dummy: 1 if fund gearing for a specific year is less than 50%, 0 if not.
  - High Gearing Dummy: 1 if fund gearing for a specific year is greater or equal to 50%, 0 if not.
- D. GAV: Fund's gross asset value for a particular year in billion euros.
- E. Fund Age: Fund age in discrete number of years, e.g.: 1-year old, 5-years old, etc.
- F. Up-Fund Dummy: 1 if fund's performance  $\geq$  0% for a particular year
- G. Fund Size Dummies:
- Small Medium Dummy: 1 if GAV  $\geq$  EUR250m but  $\leq$  EUR500m.
  - Medium Large Dummy: 1 if GAV  $\geq$  EUR500m but  $\leq$  EUR1,000m.
  - Large Dummy: 1 if GAV is  $\geq$  EUR1,000m.

To establish if there is a systematic change in the mean or variances of the data, unit root tests were performed for four key variables. Assuming individual unit root processes, both augmented Dickey-Fuller test and the Phillips-Perron test strongly reject the null hypothesis that a unit root exists, hence indicating that data is stationary and do not need to be differenced (Exhibit 9).

**Exhibit 9: Unit Root Tests on Key Variables**

Method	Fund Return	Gearing	WMR	GAV
Fischer Chi Square (Dickey-Fuller)	1699.5594 (0.000)	1056.0093 (0.000)	3319.0327 (0.000)	1236.22 (0.000)
Fischer Chi Square (Phillips-Perron)	1094.4672 (0.000)	923.7808 (0.000)	1451.387 (0.000)	2430.287 (0.000)

### Weighted Market Return (WMR)

We first test whether fund returns are driven mainly by the underlying market fundamentals, while including gearing and fund size as essential fund-level control variables. Beginning with pooled OLS estimations shown in Regression 1 of Exhibit 10, the Heteroskedasticity robust standard errors suggest that WMR is a significant driver of fund returns. For every 1% change in the WMR, fund returns change by 1.5%. Next, fund size and gearing were individually added to the model. Gearing (Regression 2) and GAV (Regression 3) were statistically significant. Overall, the pooled OLS offers strong support that the WMR, Gearing, and Fund Size are significant driver of fund performance.

**Exhibit 10: Pooled OLS of WMR, Fund Size and Gearing on Fund Return**

	(1) Fund Return	(2) Fund Return	(3) Fund Return
WMR	1.504*** (0.000)	1.526*** (0.000)	1.502*** (0.000)
Gearing		-0.130*** (0.000)	
GAV (€Bil)			0.0134*** (0.001)
C	-0.0718*** (0.000)	-0.0242*** (0.000)	-0.0797*** (0.000)
Observations (N)	1985	1640	1977
R <sup>2</sup>	0.434	0.504	0.438
Adjusted R <sup>2</sup>	0.434	0.503	0.438

Akaike Criterion	-2663.0	-2416.9	-2657.4
Schwarz Criterion	-2651.8	-2400.7	-2640.6
F-Statistic	1523.1	831.7	769.9
d.f. Model	1	2	2
d.f. Regression	1983	1637	1974
Log Likelihood	1333.5	1211.5	1331.7
Estimation	OLS	OLS	OLS

Exhibit 10: Pooled OLS of WMR, Fund Size and Gearing on Fund Return (continued)

	(5)	(6)
	Fund Return	Fund Return
WMR	1.092***	1.153***
Gearing		-0.177*
GAV (€Bil)		0.0383**
Fund Age		-0.00848***
Yr_2002	0.0152	0
Yr_2003	0.0234	0.00509
Yr_2004	0.0382*	0.0161
Yr_2005	0.0365*	0.0303
Yr_2006	0.0568***	0.0467**
Yr_2007	0.0236	0.0357**
Yr_2008	-0.0623***	-0.0351**
Yr_2009	-0.0681***	-0.0396***
Yr_2010	0.0107	0.0447***
Yr_2011	-0.0204	0.0233**
Yr_2012	-0.0471**	0
C	-0.0337*	0.0401
Observations (N)	1985	1637
R <sup>2</sup>	0.524	0.575
Adjusted R <sup>2</sup>	0.521	0.572
Akaike Criterion	-3516.9	-3132.3
Schwarz Criterion	-3449.8	-3062.1
F-Statistic	76.92	72.05
d.f. Model	11	12
d.f. Regression	344	314
Log Likelihood	1770.4	1579.1
Hausman Test	106.36	47.91
P>Chi <sup>2</sup>	(0.000)	(0.000)
Estimation	2-Way FE	2-Way FE

p-values in parentheses

\* p<0.05 \*\* p<0.01 \*\*\* p<0.001

In order to remove the effects of time-invariant characteristics from the independent variables, the fixed effects estimations were used. The WMR is a significant driver of fund returns: for every 1% increase in the WMR, fund returns increases by 1.1% (Regression 5). Gearing on average has a negative impact on fund returns (Regression 6). For a 10% increase in gearing, fund returns falls by 1.8%. In contrast, Fuerst & Matysiak (2013) reported a positive association between gearing and fund returns. That study, however, used data for 2001-2007 and did not take into account the adversely negative impact of gearing during the Global Financial Crisis that started in late 2007.

In the Fuerst & Matysiak (2013) study, it was also found that there was a negative relationship between fund size and fund returns. However, this study finds that fund performance is positively associated with fund size. For every 1bn increase in fund size, fund returns increase by 0.04%. One reason for this is a potential endogeneity issue, i.e. that better performing funds attract more

investors which then leads to larger fund sizes. Hence it may be other factors rather than fund size, which are responsible for the higher returns (Matallin-Saez, 2011).

### Gearing

Regression 6 showed that gearing on average had a negative impact on fund returns. One reason for this is that the average negative impact is due to the distortions caused by the numerous extremely high magnitude negative returns in certain funds during the Global Financial Crisis.

If gearing indeed has a negative impact on fund returns, then the question is why fund managers still employ leverage. It would also defy basic finance theory of “gearing boost returns”. Fundamentally, this is because in an up-fund situation, when fund returns are positive, leverage would boost returns. But in a down-fund situation, leverage would compound losses. As such, an alternative model was specified to capture the asymmetric effects of gearing. The estimated model is:

$$R_{(t)} = \beta_0 + \beta_1 WMR + \beta_2 Gearing + \beta_3 Dummy_1 Gearing + \beta_4 X + \varepsilon_{(t)}$$

The idea is to capture and test for asymmetric effects due to gearing exposure.  $Dummy_1$  takes the value of 1 when some fund return condition is satisfied. In this case, when the up-fund condition, defined as when fund return  $\geq 0$  is met,  $Dummy_1$  takes the value of 1, and 0 if not. Written explicitly:

$$R_{(t)} = \beta_0 + \beta_1 WMR + \beta_2 Gearing + \beta_3 UpFund \times Gearing + \beta_4 X + \varepsilon_{(t)}$$

The magnitude of negative fund returns are expected to be greater than the magnitude of positive fund returns due to the effects of interest payments. As such, the following conditions are expected to be satisfied:

- 1:  $\beta_2 \leq 0$  since in a down-fund situation, gearing’s effects are expected to be negative
- 2:  $\beta_2 + \beta_3 \geq 0$  since in an up-fund situation, gearing’s effects are expected to be positive
- From 1 and 2,  $\beta_3 \geq 0$
- $|(\beta_2 + \beta_3)| < |\beta_2|$  since the downside impact of leverage is expected to outweigh the upside impact

All three conditions are satisfied in Regression 7 (Exhibit 11). The down-fund gearing coefficient,  $\beta_2$ , suggests that when fund returns are less than 0%, for every additional 10% of gearing, fund returns are expected to decrease by 2.9% on average. The up-fund gearing coefficient,  $\beta_2 + \beta_3$ , suggests that when fund returns are greater than or equal to 0%, for each additional 10% of gearing, fund returns are expected to increase by 0.7% on average. The results are illustrated in Exhibit 12.

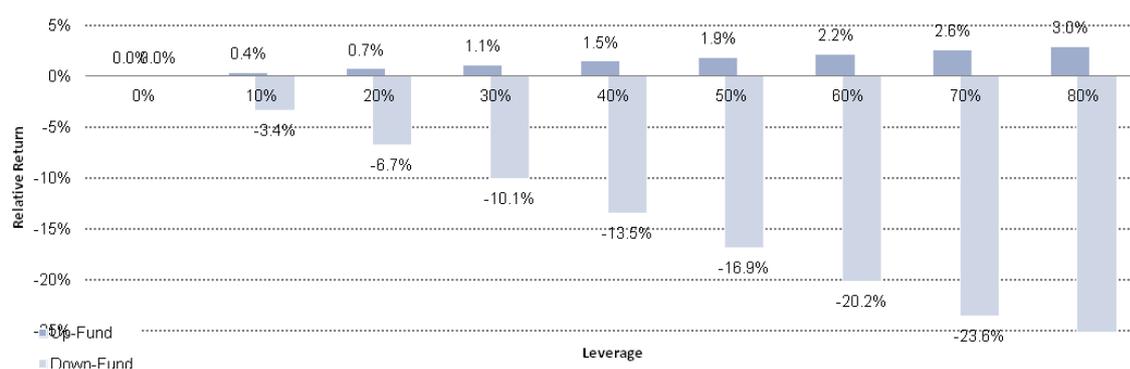
Exhibit 11: Gearing Effects

	(7)		(8)	
	Fund Return		Fund Return	
Up-Fund Gearing	0.364***	(0.000)	0.313***	(0.000)
Gearing	-0.293***	(0.000)	-0.186**	(0.003)
WMR	0.956***	(0.000)	0.967***	(0.000)
GAV (€Bil)	0.0347***	(0.000)	0.0344***	(0.000)
Fund Age	-0.00433**	(0.006)	-0.00459**	(0.003)
Up-Fund & High Gearing			0.0690*	(0.030)
High Gearing			-0.107***	(0.001)
Yr_2002	0	(.)	0	(.)
Yr_2003	0.00216	(0.876)	0.00246	(0.858)
Yr_2004	0.00327	(0.798)	0.00324	(0.799)
Yr_2005	0.00919	(0.549)	0.00886	(0.571)
Yr_2006	0.0233	(0.056)	0.0232	(0.055)
Yr_2007	0.0121	(0.233)	0.0110	(0.274)
Yr_2008	-0.00870	(0.378)	-0.0103	(0.302)
Yr_2009	-0.00870	(0.295)	-0.00792	(0.333)
Yr_2010	0.0114	(0.128)	0.0125	(0.096)
Yr_2011	0.00621	(0.349)	0.00727	(0.274)
Yr_2012	0	(.)	0	(.)
C	-0.00171	(0.946)	-0.0146	(0.558)
Observations (N)	1637		1637	
R <sup>2</sup>	0.740		0.744	
Adjusted R <sup>2</sup>	0.738		0.741	
Akaike Criterion	-3933.2		-3953.4	
Schwarz Criterion	-3857.5		-3867.0	
F-Statistic	160.9		147.1	
d.f. Model	13		15	
d.f. Regression	314		314	
Log Likelihood	1980.6		1992.7	
Hausman Test	36.91		35.86	
P>Chi <sup>2</sup>	(0.001)		(0.003)	
Estimation	2-Way FE		2-Way FE	

p-values in parentheses

\* p<0.05 \*\* p<0.01 \*\*\* p<0.001

Exhibit 12: Results of Asymmetric Gearing Effects



Next, to examine if the up-fund and down-fund gearing effects differed for highly and lowly geared funds, an alternative model was specified. The High Gearing dummy was added and interacted with the Up-Fund dummy to analyse the impact of gearing on individual fund's performance. To capture the fund size specific gearing effects, where Dummy<sub>1</sub> take the value 1 if fund

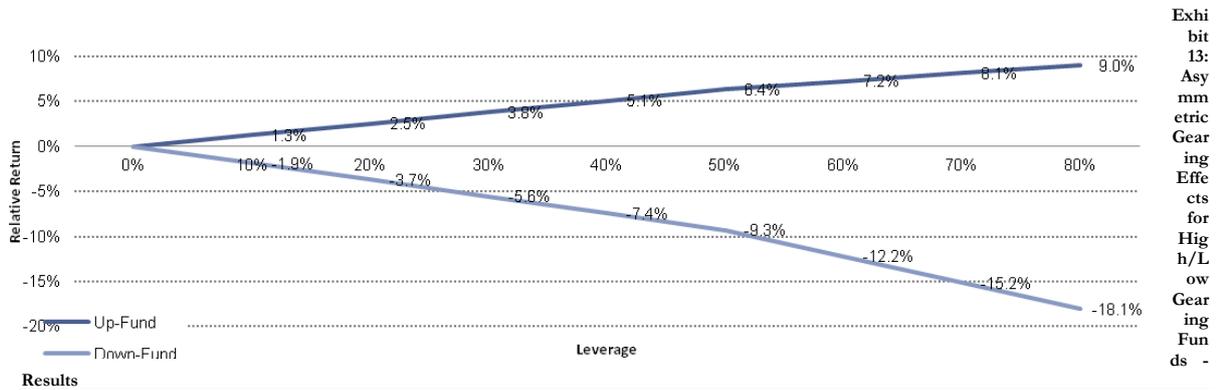
return greater or equals to 0% and 0 if not, and  $Dummy_2$  takes the value 1 if high gearing and 0 if not, the following equation was estimated:

$$R_{(t)} = \beta_0 + \beta_1 WMR + \beta_2 Gearing + \beta_3 Dummy_1 Dummy_2 \times Gearing + \beta_4 Dummy_1 \times Gearing + \beta_5 Dummy_2 \times Gearing + \beta_i X + \varepsilon_t$$

That is:

$$R_{(t)} = \beta_0 + \beta_1 WMR + \beta_2 Gearing + \beta_3 UpFund \times HighGearing \times Gearing + \beta_4 UpFund \times Gearing + \beta_5 HighGearing \times Gearing + \beta_i X + \varepsilon_t$$

From Regression 8, the benefits of taking more leverage incrementally tapers off. In an up-fund situation, for low gearing funds, every additional 10% leverage results in +1.3% fund returns on average (as calculated by  $\beta_2 + \beta_3$ ). However, when gearing exceeds 50%, for every additional 10% in leverage, fund returns are only expected to increase by 0.9% on average (as calculated by  $\beta_2 + \beta_3 + \beta_4 + \beta_5$ ). The increased risks associated with higher debt levels would have likely increased the costs of borrowing, which offsets a large part of the earnings enhancing capacity of leverage. The results are illustrated in Exhibit 13.



In a down-fund situation, when a fund is highly geared, the marginal loss is amplified. For low gearing funds, every additional 10% of leverage results in an expected fund return of -1.9% on average (coefficient value of  $\beta_2$ ). And for funds with gearing in excess of 50%, every additional 10% of leverage results in fund return of -2.9% on average (as calculated by  $\beta_2 + \beta_5$ ). This is largely because, in addition to subpar asset level performance, funds will still have to service their debt hence contributing to further earnings erosion.

One caveat is that in addition to interest payments, fixed management fees and variable performance fees could have been the cause of the observed asymmetric effects. For example, management fees may typically take the 2/20 structure, meaning that a base management fee equivalent to 2% of commitment capital is charged, and an additional 20% of excess returns is charged as

performance fee (Moszoro, 2012). If the partnership agreement (the Limited Partner being the investor) sets a target return of 8%, and if the fund delivers 15%, then 20% of the 7% outperformance, (which equates to 1.4% of returns) is the performance fee.

In the down-fund situation, only the flat base management fee is charged and there is no further fund return leakage in the form of performance fees. Hence, fund returns leakage due to fees is independent of fund gearing levels. As such, the observed incremental loss of fund returns with high gearing can safely be attributed to the effects of gearing. However, in an up-fund situation, a transmission mechanism of higher gearing and higher performance fees and hence fund return leakage exists. Higher gearing boosts fund returns and hence there will be higher performance fee leakage. Therefore, the lower marginal benefits of gearing with higher gearing in an up-fund situation could also be due to the effect of performance fees and is something to be investigated if future available data includes fund fee information.

### Macroeconomic Drivers & Competing Asset Classes

Finally, the research examines the potential statistical relationships between fund returns of NLREFs and that of other competing asset classes including EU Bond Yields, EU Stock Returns and Listed Real Estate Vehicles (EU REITs). The foreign exchange effects on fund returns were also examined with a weighted FX variable, constructed in the same way as the WMR. The EU Gross Domestic Product growth was included in the models to serve as a control for the impacts of macro-economic growth on all asset classes.

As expected, EU GDP growth, EU Bond Yields, and EU stock returns were positively associated with and are significant drivers of NLREF returns. As long-term bond yields are somewhat a proxy for risk free cost of capital, it is plausible that the higher risk investments such as NLREFs' performance displays certain correlation with interest rates through several linkages such as cost of capital.

With REITs and non-listed fund returns driven largely by their underlying property assets, REITs have exhibited the expected positive and significant relationship with NLREF returns (Regression 10). For every 1% increase in EU REIT returns, NLREFs returns are 0.1% on average. The coefficient magnitude and sign are aligned with research findings that REITs are positively correlated with equities, but are more volatile. Finally, the Weighted FX was also a significant driver of fund returns, with an unfavourable impact on average.

Exhibit 14: Competing asset classes

	(9)		(10)		(11)	
	Fund Return		Fund Return		Fund Return	
Gearing	-0.145***	(0.000)	-0.215*	(0.016)	-0.158	(0.068)
Fund Age	-0.00163***	(0.001)	-0.0179***	(0.000)	-0.0163***	(0.000)
GAV (€Bil)	0.0146**	(0.003)	0.0325*	(0.014)	0.0198	(0.077)
EU GDP	1.053***	(0.000)	0.528*	(0.021)	0.391	(0.121)
EU Bonds	7.055*	(0.011)				
EU Stocks	0.141*	(0.022)				
EU REITS			0.114***	(0.000)		
Weighted FX					-0.962***	(0.000)
Yr_2002	-0.0489	(0.489)	0	(.)	0	(.)
Yr_2003	-0.0491	(0.493)	0.00650	(0.717)	0.0612**	(0.002)
Yr_2004	-0.0309	(0.674)	0.0229	(0.216)	0.0483**	(0.001)
Yr_2005	0.0303	(0.624)	0.0823***	(0.000)	0.102***	(0.000)
Yr_2006	0.0231	(0.748)	0.0871***	(0.000)	0.132***	(0.000)
Yr_2007	-0.0925	(0.212)	0.0840***	(0.000)	0.0422**	(0.007)
Yr_2008	-0.168***	(0.000)	-0.0637**	(0.007)	-0.0938***	(0.000)
Yr_2009	-0.126**	(0.002)	-0.0607***	(0.000)	-0.0133	(0.507)
Yr_2010	0.00945	(0.790)	0.0826***	(0.000)	0.0691***	(0.000)
Yr_2011	0.0152	(0.428)	0.0878***	(0.000)	0.0589***	(0.000)
Yr_2012	0	(.)	0	(.)	0	(.)
C	-0.135*	(0.021)	0.151***	(0.000)	0.142***	(0.000)
Observations (N)	1597		1597		1595	
R <sup>2</sup>			0.426		0.472	
Adjusted R <sup>2</sup>			0.421		0.467	
Akaike Criterion	.		-2568.9		-2695.5	
Schwarz Criterion	.		-2493.7		-2620.2	
F-Statistic			33.59		47.55	
d.f. Model	16		13		13	
d.f. Regression			308		308	
Log Likelihood			1298.5		1361.7	
Chi <sup>2</sup>	516.9					
Hausman Test	16.07		53.78		1361.7	
P>Chi <sup>2</sup>	(0.139)		(0.000)		(0.001)	
Estimation	RE		2-Way FE		2-Way FE	

p-values in parentheses

\* p&lt;0.05 \*\* p&lt;0.01 \*\*\* p&lt;0.001

## 4 Conclusions and further work

Employing fixed and random effect panel data models, this paper examined a large number of non-listed real estate funds based on the 2001-2012 INREV database. The research studied how fund size, gearing, WMR, competing asset classes and economic growth influenced fund performance.

The underlying market sector and country returns, as measured by the WMR, contributed most significantly to fund returns. Gearing, fund size, competing asset classes and GDP growth were also significant drivers of fund returns.

In contrast to previous studies, larger funds, on average, performed better and further work is required to identify the reasons for this. In particular, is this due to the diversification benefits, economies of scale or the ability to undertake large investments where fewer competitors exist? As previous studies show, liquidity was also an important factor in the fund size-performance argument so future research may consider evaluating its effects on NLREFs. It was also found that the incremental fund size effects tapered off. Medium-large funds performed only marginally better than large funds, on average. This raises the question whether it is possible to determine optimum size for a non-listed real estate fund based on empirical performance data.

As the data on so-called opportunistic funds were not available for this study, the research findings do not apply to the entire NLREF universe, being limited only to core and value-added funds. Also, as data on opportunistic funds becomes available together with a larger sample size, it will be possible to perform a more comprehensive study on the impact of fund performance drivers by investment style and investment strategy.

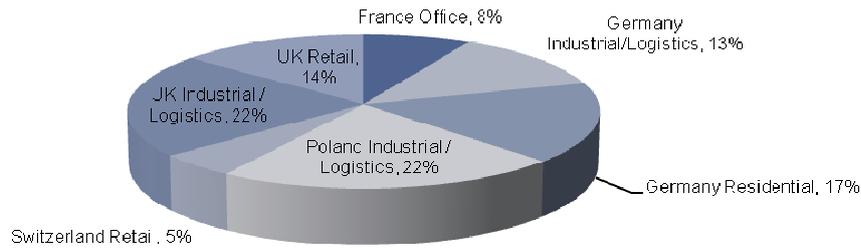
The underlying market sector and country returns together with fund characteristics and external factors, which were analysed in this research, explain more than half of the variation in performance of non-listed real estate funds. Other factors need to be identified and analysed in order to see if a better understanding of fund performance is possible. A possible extension of this study is to analyse if fund managers are able to generate alpha, despite a number of known problems in the empirical measurement process of fund.

Real estate risk has not been explicitly addressed in the analysis and which is necessary in order to have a complete picture of the underlying performance. From this perspective, aspects which need to be taken into consideration include: the accuracy of real estate valuations in different market/property cycle environments; diversification considerations, including the underlying correlations across property types and countries. More broadly, these, together with property's correlations with other assets classes, become important when looking at the risk of banks' balance sheets, where lending on real estate is involved.

## Appendix

Exhibit 15 shows how the WMR was constructed for a particular fund in 2011. Using a weighted-average of their allocation and the Investment Property Databank (IPD) returns for the specific country-sector, the weighted return is the WMR.

**Exhibit 15: Example of WMR construction**



Year	Country	Sector	Allocation	IPD Country-Sector Index Return	Weighted Return
2011	France	Office	8.00%	7.25%	0.58%
2011	Germany	Industrial / Logistics	12.50%	6.99%	0.87%
2011	Germany	Residential	16.80%	7.84%	1.32%
2011	Poland	Industrial / Logistics	21.80%	6.36%	1.39%
2011	Switzerland	Retail	5.00%	8.50%	0.43%
2011	United Kingdom	Industrial / Logistics	21.80%	7.26%	1.58%
2011	United Kingdom	Retail	14.10%	7.10%	1.00%
	<b>WMR</b>		<b>100%</b>		<b>7.17%</b>

## References

- Aghion, P., & Tirole, J. (1997). Formal and real authority in organizations. *Journal of political economy*, 1-29.
- AltAssets. (2006). *A discussion on the J-Curve in private equity*. Retrieved July 20, 2013, from AltAssets.com: A discussion on the J-Curve in private equity
- Amoako, B., & Samarbakhsh, A. L. (2012). The role of leverage in hedge fund failure during the 2008 financial crisis. *Journal of Finance*, 345-678.
- Baltagi, B. H. (2008). *Econometric Analysis of Panel Data* (4 ed.). New York: John Wiley and Sons Ltd.
- Barry, P. (1980). Capital asset pricing and farm real estate. *American Journal of Agricultural Economics*, 62, 549-553.
- Baum, A. (2009). *Commercial Real Estate Investment: A Strategic Approach*. London: Estates Gazette.
- Baum, A., & Farrelly, K. (2009). Sources of alpha and beta in property funds: a case study. *Journal of European Real Estate Research*, 218-234.
- Black, F., Jensen, M. C., & Scholes, M. (1972). *The Capital Asset Pricing Model: Some Empirical Tests*. New York: Praeger Publishers.
- Breidenbach, M., & et al. (2006). Determining real estate betas for markets and property types to set better investment hurdle rates. *Journal of Real Estate Portfolio Management*, 12, 73-80.
- Brounen, D., Veld, H. O., & Raitio, V. (2007). Transparency in the European non-listed real estate funds market. *Journal of Real Estate Portfolio Management*, 13, 107-118.
- Brown, R., & Matysiak, G. (2000). *Real Estate Investment: A Capital Market Approach*.
- Busse, J. A., Goyal, A., & Wahal, S. (2010). Performance and persistence in institutional investment management. (T. J. Finance, Ed.) *The Journal of Finance*, 65, 765-90.
- Charles, A., & Darne, O. (2005). Outliers and GARCH models in financial data, *Economics Letters*, 86, 347-52.
- Chen, J., Hong, H., & Kubik, M. H. (2004). Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization. *The American Economic Review*, 94(5), 1276-1302.
- Chong, J., Halcousis, D., & Phillips, M. G. (2012). Misleading Betas: An Educational Example. *American Journal Of Business Education*, 618.
- Deangelo, H., & Masulis, R. (1980). Optimal capital structure under corporate and personal taxation. *Journal of Financial Economics*, 8, 3-29.
- Deloitte Touche Tohmatsu. (2010). *Review of reporting best practice for non-listed real estate vehicles in Asia*.
- Devaney, S., Lee, S., & Young, M. (2007). Serial persistence in individual real estate returns in the UK. *Journal of Property Investment and Finance*, 25, 241-73.
- Dietz, P. (1966). *Pension Funds: Measuring Investment Performance*. Free Press.

- Diller, C., Heger, I., & Wulff, M. (2009, Apr). *The private equity J-Curve: Cash flow considerations from primary and secondary points of view*. Retrieved July 2013, from Capital Dynamics: [http://www.capdyn.com/newswriter\\_files/pe-mathematics-apr2009-private-equity-j-curve.pdf](http://www.capdyn.com/newswriter_files/pe-mathematics-apr2009-private-equity-j-curve.pdf)
- Fama, E. F., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33, 3–56.
- Franzoni, Nowak, Francesco, & Phalippou, L. E. (2012). Private Equity Performance and Liquidity Risk. *Journal of Finance*, 67(6), 2341-2373.
- Frecka, T. F., & Hopwood, W. S. (1983). The effects of outliers on the cross-sectional distributional properties of financial ratios. *The Accounting Review*, 58, 115–28.
- Fredman, A., & Wiles, R. (1993). How mutual funds work. *New York Institute of Finance*.
- Fuerst, F., & Marcato, G. (2009). Style analysis in real estate markets: beyond the sector and region dichotomy. *The Journal of Portfolio Management*, 35, 104–17.
- Fuerst, F., & Matysiak, G. (2013). Analysing the performance of nonlisted real estate funds: a panal data analysis. *Applied Economics*, 1777-1788.
- Georgiev, G., Gupta, B., & Kunkel, T. (2003). Benefits of Real Estate Investment. *The Journal of Portfolio Management*, Special Issue.
- Gottschalg, O., Talmor, E., & Vasvari, F. (2010). Private Equity Fund Level Return Attribution: Evidence from U.K. Based Buyout Funds. *Journal of Finance*.
- Grabenwarter, U. (2005). *Exposed to the J-Curve: Understanding and Managing Private Equity Fund Investments*. Retrieved July 21, 2013, from [http://www.carriedinterest.com/2006/02/weekly\\_private\\_\\_2.html](http://www.carriedinterest.com/2006/02/weekly_private__2.html)
- Grauer, R., & Hakansson, N. (1995). Gains From Diversifying Into Real Estate: Three Decades of Portfolio Returns Based on the Dynamic Investment Model. *Real Estate Economics*, Vol.23.
- Grossman, S., & Hart, O. (1982). Corporate financial structure and managerial incentives. The economics of information and uncertainty. *University of Chicago Press*.
- Gruber, M. J. (1996). Another puzzle: the growth in actively managed mutual funds. *Journal of Finance*, 51, 783–810.
- Haran, M., McGreal, W. S., Adair, A. S., & Webb, J. R. (2008). Non-listed property fund investment in urban renewal/ regeneration property in the UK: the impact of investment style on decision making. *Journal of Real Estate Portfolio Management*, 14, 185–93.
- Hastings, C. M. (1947). Low Moments for Small Samples: A Comparative Study of Order Statistics. *Annals of Mathematical Statistics*, 413-426.
- Hellmann, T., & Puri, M. (2002). Venture capital and the professionalization of Start-up firms: Empirical evidence. *Journal of Finance*, 57, 169-197.
- Hoesli, C., & Lekander, J. (2005). Suggested versus actual institutional allocations to real estate in Europe: matter of size? *Journal of Alternative Investments*, 62-70.
- Hsiao, C. (2003). *Analysis of Panel Data* (2nd ed.). New York: Cambridge University Press.
- INREV. (2008). *INREV guidelines*. Amsterdam.
- INREV. (2011). *Vehicles database analysis*. Amsterdam.

- Jensen, M. (1968). The Performance of Mutual Funds in the Period 1945-1964. *Journal of Finance*, 389-416.
- Kallberg, J., & Liu, C. (1996). The Role of Real Estate in the Portfolio Allocation Process. *Real Estate Economics*, Vol. 24.
- Key, T., & Lee, S. (2008). *Real estate fund styles*. Report for INREV.
- Koop, G., Poirier, D. J., & Tobias, J. (2005). Bayesian Semiparametric Inference in Multiple Equation Models. *Journal of Applied Econometrics*, 20, 723-747.
- Kraus, A., & Litzenberger, R. (1973). A state - Preference model of optimal financial leverage. *The Journal of Finance*, 911-922.
- Latzko, D. (1999). Economies of scale in mutual fund administration. *Journal of Financial Research*, 22, 331-340.
- Lee, S. (2003). The persistence of real estate fund performance. *American Real Estate Society Meeting*, Monterey, CA.
- Lee, S., Young, M. S., & Devaney, S. (2007). Non-normal real estate return distributions by property type in the UK. *Journal of Property Research*, 23, 109-33.
- Lerner, J., Hardyman, F., & Leamon, A. (2004). *Venture capital and private equity: A casebook* (3rd ed.). John Wiley & Sons.
- Lim, W. (2013). Property in Diversified Portfolios: A UK Perspective. *Working Paper*.
- Ling, D. C. (2005). A random walk down main street: can experts predict returns on commercial real estate?., *Journal of Real Estate Research*, 27, 137-54.
- Lizieri, C., & Satchell, S. (1997). Interactions between property and equity markets: An investigation of linkages in the United Kingdom 1972-1992. *The Journal of Real Estate Finance and Economics*, 15, 11-26.
- Lowenstein, R. (1997). Frightened funds: Is there a master in the house. *Wall Street Journal*, 20, C1.
- Lu, K., & Mei, J. (1999). The return distributions of property shares in emerging markets. *Journal of Real Estate Portfolio Management*, 5, 145-160.
- Lynch, A., & Musto, D. (2003). How investors interpret past fund returns . *The Journal of Finance*, 58, 2033-58.
- Marcal, E. F., Pereira, P. L., Martin, D. M., & Nakamura, W. T. (2011). Evaluation of contagion or interdependence in the financial crises of Asia and Latin America, considering the macroeconomic fundamentals. *Applied Economics*, 43, 2365-79.
- Mark, C., & Asieh, M. (2005). Real estate in the real world: Dealing with non-normality and risk in an asset allocation model. *Journal of Real Estate Portfolio Management*, 11, 37-53.
- Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 77-91.
- Matallin-Saez, J. C. (2011). On causality in the size-efficiency relationship: the effect of investor cash flows on the mutual fund industry. *Applied Economics*, 43, 4069-79.
- Mei, J., & Lee, A. (1994). Is there a real estate factor premium? *The Journal of Real Estate Finance and Economics*, 9, 113-126.
- Mei, J., & Liu, C. H. (1994). The predictability of real estate returns and market timing. *The Journal of Real Estate Finance and Economics*, 8, 115-35.
- Milgrom, P., & Roberts, J. (1988). An economic approach to influence activities in organizations. *American Journal of Sociology*, S154-S179.

- Modigliani, F., & Miller, M. (1963). Corporate income taxes and the cost of capital: a correction. *The American Economic Review*, 53, 433-443.
- Moszoro, M. (2012). The Math of Private Equity: 2/20, 3-on-5, and 1-in-8. *Working Paper*.
- Okunev, J., Wilson, P., & Zurbrugg, R. (2000). The causal relationship between real estate and stock markets. *The Journal of Real Estate Finance and Economics*, 21, 251-61.
- Perold, A., & Salamon JR, R. (1991). The right amount of assets under management. *Financial analysts journal*, 31-39.
- Pin, S. Y., Lu, C., & Lin, Z. (2012). International Real Estate Mutual Fund Performance: Diversification or Costly Information? *Journal of Real Estate Finance & Economics*, 44(3), 394-413.
- Quan, D. C., & Titman, S. (1999). Do real estate prices and stock prices move together? An international analysis. *Real Estate Economics*, 27, 183-207.
- Rom, B. (1993). Post-Modern Portfolio Theory Comes of Age. *Journal of Investing*.
- Schluter, C., & Trede, M. (2008). Identifying multiple outliers in heavy-tailed distributions with an application to market crashes, *Journal of Empirical Finance*. *Journal of Empirical Finance*, 15, 700-13.
- Schneeweis, T., Martin, G., Kazemi, H., & Karavas, V. (2005). The Impact of Leverage on Hedge Fund Risk and Return. *Journal of Alternative Investments*, 7(4), 10-21.
- Sharpe, W. (1964). Capital Asset Prices: A theory of Market Equilibrium under conditions of risk. *Journal of Finance*, 19, 425-442.
- Sharpe, W. (1994). The Sharpe Ratio. *The Journal of Portfolio Management*, 49-58.
- Stevenson, S. (2000). International real estate diversification: empirical tests using hedged indices. *Journal of Real Estate Research*, 19, 105-131.
- Stevenson, S. (2006). Driving forces in the performance attributes of real estate funds. *INREV Conference*. Rome.
- Sun, B. H., & Hsiao, C. (2000). *To pool or not to pool panel data, in Panel Data Econometrics : Future Directions*. (J. Krishnakumar, & E. Ronchetti, Eds.) Amsterdam, North Holland.
- Treynor, J. (1961). Toward a Theory of Market Value of Risky Assets. *Unpublished manuscript*.
- Williamson, O. (1975). Markets and hierarchies: antitrust analysis and implications. *New York: The Free Press*.
- Williamson, O. (1988). Corporate finance and corporate governance. *The Journal of Finance*, 43, 567-591.
- Yan, X. (2008). Liquidity, Investment Style, and the Relation between Fund Size and Fund Performance. *Journal of Financial & Quantitative Analysis*, 43(3), 741-767.
- Yang, S.-F., & Wu, T.-H. (2012). More Evidence on the Fund Performance and Fund Size: A Semi-Parametric Smooth-Coefficient Approach. *International Research Journal of Finance & Economics*(101), 165-168.
- Young, M. S., & Graff, R. A. (1996). Systematic behavior in real estate investment risk: performance persistence in NCREIF returns. *Journal of Real Estate Research*, 12, 369-81.



---

## 14. The development of commercial property prices and values analysis

Wojciech Doliński



Wojciech Doliński – MRICS

**Acknowledgments:**

I would like to take this opportunity to express my gratitude and deep regards to Christopher Grzesik FRICS, who introduced me to the valuation profession and provided me with guidance and gave me an opportunity to discuss real estate market analysis. I would like to thank colleagues working in the real estate sector for exchange of ideas and comments received during the Workshop *Current Trends in Real Estate*, organised by Narodowy Bank Polski in November 2013.

I also thank Joe Mannina Jr., Chief Operating Officer of RCA Inc. for permission to include data on international property markets in this paper.

---

## Article 14 Contents

Abstract .....	100
Introduction .....	101
1. Primary data .....	102
Data availability constraints .....	102
2. Data providers .....	105
The beginning .....	105
3. Investment flow and data exchange .....	107
4. Post-crisis reality .....	109
Market indicators evolution and new data products .....	109
5. Market data analysis and presentation .....	112
What is the yield? .....	113
Spatial data .....	114
In search of local indicators .....	116
6. Conclusion .....	119
References .....	120
Annex 1 .....	121
Annex 2 .....	124

## ABSTRACT

---

Presented paper provides a general overview of the commercial real estate market analysis both internationally and locally, aiming to emphasize the importance of coherent market analysis and indicators in today's international investment market.

Unfortunately, the commercial real estate market lacks coherent information and limits market transparency. The availability of data enables to suggest a certain market division, varying from undeveloped non-transparent markets through developed markets with limited market data to highly transparent developed markets with price indexes.

The responsibility for the task of developing market analysis of the commercial real estate sector was assumed by the private sector incl. IPD databank, research departments in big consultancies and finally CoStar and Real Capital Analytics taking over a leading role in commercial property data analysis.

From the beginning of XXI century interest of market analysis shifted from pure price monitoring into researching wider market sentiments. New data products have emerged not only on the wave of a developing information economy, but also as an answer to the post-crisis global real estate reality.

However, in the circumstances of lack of market indicators coupled with market uncertainty, local consultants such as valuers have to resort to various forms of benchmarking their results, including reinstatement costs or land-to-capital ratios.

***Keywords: Commercial real estate, market analysis, yields, market transparency***

***JEL Classification: L85, R31, R33, L860, R39***

---

## **Introduction**

It has been over hundred years since a famous real estate appraiser Homer Hoyt first analysed property cycles. However, indicators for monitoring commercial real estate markets are still limited. Even nowadays many less developed markets lack primary data representing for each of the commercial sectors.

Due to the fact that basic information on prices lies at the fundamentals of further analysis it limits market transparency.

## **1. Primary data**

Analysis of primary data provides the backbone for decision making process in more efficient capital markets. Even in Poland the residential real estate market is already subject to comprehensive analysis, whereas the commercial real estate market lacks coherent information, both locally and internationally. Commercial property research ought to be more sophisticated with more attention paid to facets, stemming from primary data collection.

### **Data availability constraints**

Access to information and further analysis is determined by limitless factors including legal and local characteristics, but first and foremost – the availability of data.

This factor is notably visible on the Russian market where no commercial transactions were recorded or defined as arms-length transactions until 2000. Data concerning commercial real estate transactions located outside of Moscow are virtually unrecorded. This is clearly evident from the fact that Real Capital Analytics, Inc. (RCA), the leading data analyst, only has transaction analysis dating back to 2007.

As a result, lack of transparency deters many international investors and banks from operating in the buoyant Russian real estate market.

Table 1: Top commercial real estate investors in Russia

Russia					
Top Buyers (Prior 24 months)					
Rank▲	Buyer (click for profile)	Location	Capital Group	Acq in mil.	# Props
1	O1 Properties	Moscow/ Russia	Private	€655.9	4
2	VTB Bank	Saint Petersburg/ Russia	Institutional	€483.6	3
3	PSN Group	Moscow/ Russia	Private	€480.6	2
4	TPG Capital	Fort Worth, TX/ US	Equity Fund	€401.1	1
5	UFG Asset Management	Moscow/ Russia	Equity Fund	€266.9	2
6	Sberbank	Moscow/ Russia	Institutional	€261.4	2
7	Bin Group	Moscow/ Russia	Private	€232.4	1
8	Stroygazconsulting	Moscow/ Russia	Private	€162.5	1
9	Alfa Group	Moscow/ Russia	Institutional	€165.6	2
10	Heitman	Chicago, IL/ US	Institutional	€132.7	1
11	BIN Bank	Moscow/ Russia	Institutional	€140.1	1
12	Otkritie	Moscow/ Russia	Institutional	€93.5	1
13	Hines	Houston, TX/ US	Institutional	€66.8	1
14	Raiffeisen Bank	Wien, Vienna/ Austria	Institutional	€60.4	1
15	Kudla Group	Moscow/ Russia	Private	€64.8	1
16	Eastern Property Holdings	Geneva/ Switzerland	Public	€53.9	1
17	Svyaz Bank	Moscow/ Russia	Institutional	€39.7	1
18	Business Alliance	Moscow/ Russia	Private	€36.1	1
19	Sponda Plc	Helsinki/ Finland	Public	€38.1	1
20	Unicredit Group	Milan, Lombardy/ Italy	Institutional	€29.3	1

Source: Real Capital Analytics

There are also legal constraints, limiting access to primary data, as in the case of Poland and Greece, where access to public records is only given to certified valuers or tax authorities. Moreover, there are strict limitations on joining such professional groups, which severely limits market transparency.

In many developing countries, it is related to the everyday functioning of the particular real estate consultancy market, as well as the market segmentation in such countries.

It is very often case that the global consultancy companies operate in the vast majority of the market and use their privileged position to limit the availability of investment data to the market players. This is quite evident in Brazil, where information outside major markets can only be obtained through close relationships with global consultancies.

Different tenure regulations has an influence on every market analysis, which is quite evident in the Chinese market where ownership of land is non-existent and land leases are granted to private individuals or legal entities for periods between 30 to 50 years. What is more, the supply side of investment products is limited for overseas investors as no foreign entity can own a building within 500m of any public office.

A local example from the Polish market is the right of perpetual usufruct which very often leads to overpriced land tenders from the state treasury. The prospective user will have to pay an initial payment of 15-25%, and annual fees are levied again in service charges to future tenants. As a consequence, development land in perpetual usufruct very often exceeds the price charged for ownership rights, not because of the location but because of the illogical local law. Therefore it is important to be diligent while researching the Polish market of development land.

While researching Asian markets, special attention has to be given to built-up intensity or the impact of numerology which is very often has a significant influence on pricing and values.

The available market evidences, suggests that properties of similar characteristics with, technically, the same yielding potential, receive different sale prices, owing in the main favourable administrative numbers. And population density and built-up potential creates a situation in which land constitutes as much as 80% of the value of the completed development.

Depending on data availability and market analysis, a certain market division can be proposed, varying from undeveloped non-transparent markets through developed markets with limited market data to highly transparent developed markets with reliable price indexes.

---

## 2. Data providers

In line with market development, market analysis becomes an intrinsic component of every investment process and a vital part of every report. Rising demand has created different types of research, ranging from sector analysis, databases, strategic and portfolio analysis up to prognosis.

Even in most developed countries, monitoring of the commercial real estate market only started relatively recently in the 1970s. It goes without saying, it is still barely existent in some less developed markets like the Ukraine or the Balkans.

### The beginning

It is worth noting that in the United Kingdom the Department of the Environment had been gathering data on new construction orders, compiling them into overall stock statistics for each sector. Other official sources of relevant data such as the Central Statistical Office. However, none of these initiatives have managed to meet the increased demand coming from the booming modern commercial sector.

Owing to the fact that the public sector was reluctant to develop market analysis of the commercial of the real estate sector, the responsibility for the task was assumed by the private sector. It was distinctively visible during the 1980s in the United Kingdom, where IPD databank and small research departments in big consultancies like Cushman and Wakefield and CBRE started to appear. However, it was not long until CoStar and RCA taken over a leading role in commercial property data analysis.

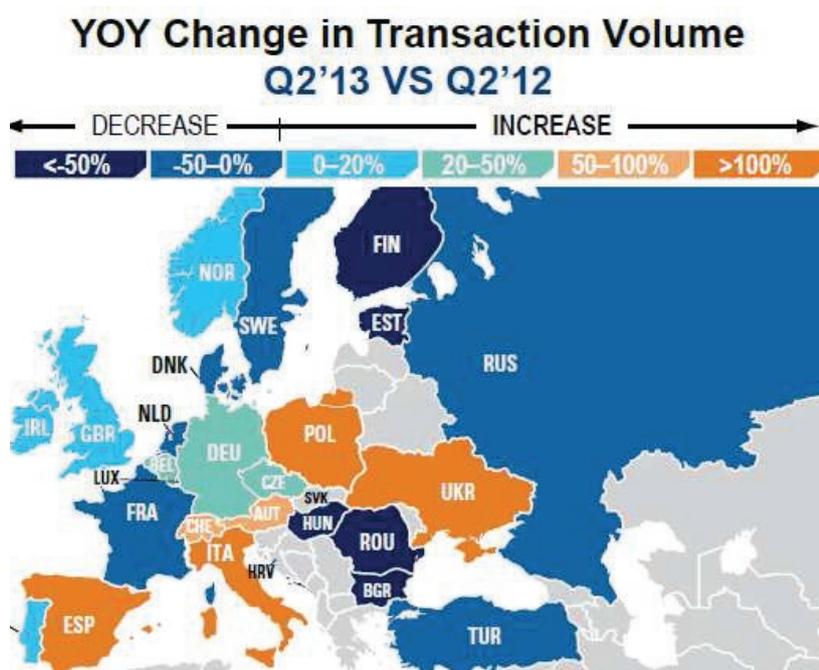
An increase in regulations and the maturing of other markets, resulted in the development of real estate markets across Europe. The Anglo-Saxon model of analysis spread not only across many Central and Eastern

European countries, but also on other continents. This sometimes occurred as a result of pre-existing colonial ties, which helped to consolidate influence in active business areas.

### 3. Investment flow and data exchange

An increasing level of cross border investment activity performed by international funds has brought the need for closer monitoring of developing markets in CEE and CIS. The map below demonstrates that the top markets in Europe witnessed declines in transaction rates and highlights the contribution of secondary markets, where market information is scarce.

Graph 1: Map depicting altering investment activity in Europe



Source: Real Capital Analytics

This trend has produced the need to find a universal way of reporting and exchanging market information. One of the organizations which is trying to address these issues is Open Standard Consortium for Real Estate (OSCRE). It aims to standardize data gathering models to enable real estate related information to be exchanged more effectively.

One of the aims of the organization is to enable an exchange of real-estate-related information to reduce data risk for market stakeholders

such as pension funds, asset management and consultancies with reduced data risk.

But, first and foremost, the aim is to provide a faster, seamless data flow without the need for data editing. Thereby, accelerating decision-making process. In the same time, it would enable investors to shorten data time lag evidently visible on the real estate market.

However, there is still a vast amount of primary data missing on the developing markets to collate "big data."

---

## 4. Post-crisis reality

Despite the fact that many mature markets have developed detailed indexes and cyclical analysis, none of these tools prevented economies from encountering economies that were suffering from a serious slump in the real estate market. There were several national examples where the cyclical nature of real estate markets was related more to the intrinsic issues than a wider international credit crunch, let alone issues rising from September 2008 and the collapse of Lehman Brothers.

Is it true that five years after the worst financial crisis in 80 years, markets are better monitored, analysed and as a result safer?

Until 2008 all the commercial analysis were concentrated on sales data. There has not been much analysis of financing component of the equation nor distressed sales monitoring, which is a corner-stone of activity in the commercial real estate market.

### **Market indicators evolution and new data products**

So what has changed from the beginning of XXI century in respect of market analysis?

As already mentioned, interest shifted from pure price monitoring into researching wider market sentiments. As an example one might look at troubled asset sales analysis compound by Real Capital Analytics.

Table 2: Distressed assets, list of German properties

OFFICE | Europe Most recent 100 results (as of 10/13/2013)

PROPERTY TRADES TROUBLED ASSETS

MAP

PLOT COMP SET CHANGE TO 1-LINE VIEW

Plot These	Status Date	Troubled Subtype	Type/ Subtype	Property Name Address City,State/Country	m <sup>2</sup> Yr. Bldg/Renov # bldg/flrs	Owner/Buyer Lender	Comments
<input type="checkbox"/>	Troubled Feb'13	Stalled/Incomplete; Owner/GP Bankrupt; Receivership, Admin, Special	Office - CBD	City Deco Na Strzi And Budejovicka Prague/ CZ	17,000 m <sup>2</sup> 2014 1 bldg/6 flrs	ECM Real Estate Development	
<input type="checkbox"/>	Troubled Aug'13	Delinquent/Default; Owner Financially Challenged	Office - CBD	The Sqaire B43 Frankfurt, Hesse/ DE	145,000 m <sup>2</sup> 2011 1 bldg/9 flrs	IVG ImmobilienJV	
<input type="checkbox"/>	Troubled Jun'13	Delinquent/Default; Foreclosure Initiated; Maturity Default/Past Due; Owner Financially Challenged; Receivership, Admin, Special	Office - CBD	Berlin, Frankfurt & Hamburg Office DE	434,580 m <sup>2</sup>	Arminius Funds Mgmt	Last Traded May-12
<input type="checkbox"/>	Troubled Jan'12	Maturity Default/Past Due; Transferred to Special Servicer	Office - Sub	Essen Lindenallee 6-8 Essen, North Rhine-Westphalia/ DE	11,558 m <sup>2</sup> 1 bldg/9 flrs	Cerberus CMBS: Hatfield Philips OBO TMAN 7	Last refinanced Apr-07 (\$24.3 mil); 1st mtg: \$19.0 mil (alloc'd.)
<input type="checkbox"/>	Troubled Jan'12	Maturity Default/Past Due; Transferred to Special Servicer	Office - CBD	Hamburg Alte Holstenstrasse 44-46 Hamburg, Hamburg/ DE	11,029 m <sup>2</sup> 6 flrs	Cerberus CMBS: Hatfield Philips OBO TMAN 7	Last refinanced Apr-07 (\$24.3 mil); 1st mtg: \$19.0 mil (alloc'd.)
<input type="checkbox"/>	Troubled Jan'12	Maturity Default/Past Due; Transferred to Special Servicer	Office - CBD	Hanover 1 Rathenaustrasse 3 Hanover, Lower Saxony/ DE	4,407 m <sup>2</sup> 1 bldg/6 flrs	Cerberus CMBS: Hatfield Philips OBO TMAN 7	Last refinanced Apr-07 (\$10.4 mil); 1st mtg: \$8.1 mil (alloc'd.)

Source: Real Capital Analytics

The global data provider and analytics firm, RCA, has not only expanded their research programme into such areas as distressed transactions, but has also developed new products like Troubled Assets Radar and TrendTracker™ (Example see Annex 2). Identifying troubled assets is facilitated by monitoring bankruptcy announcements, court administration, tenant defaults or debt restructuring.

Moreover, in May 2012 it launched a co-branded market index with renowned rating agency Moody's, to measure the change in commercial property prices in the United States. It is based on a repeat-sales regression methodology, which is derived from transaction evidence. It appears to be an objective measure of the commercial property prices for markets in the US and more recently, the United Kingdom.

It worth noting that, this kind of transaction-based index proved to be a relevant and efficient measurement mechanism in the large US market which saw transactions totalling over USD 30 billion. For smaller

markets, unfortunately, building an index on the basis of repeated sales would not prove to be statistically relevant.

It is quite evident that described above data products have emerged not only on the wave of a developing information economy, but also as an answer to the post-crisis global real estate reality.

Most of the less developed markets cannot boast the luxury of market indices or reliable indicators like TrendTracker TM, described above.

Therefore, consultants and analysts have to resort to other types of benchmarking.

Following the credit crunch, many adjustments have also been made on the consultancy side of market operations. Both national and international standards for valuers have been amended, putting more emphasis on valuation uncertainty, market risks and prognosis (Whatling 2007).

The real-estate-related aspects of the credit-crunch have shown that valuation is one of the cornerstones of financing decisions.

## 5. Market data analysis and presentation

Despite the fact that many mature markets can boast wider market information, most stakeholders in developing economies have to rely on pure sale price data. Despite the fact that, generally Central and Eastern European countries used to be considered as non-transparent, several years ago there seemed to be a visible difference between the countries. It can be exemplified by changing rankings of Poland and Czech Republic within Jones Lang La Salle transparency index. This is widely caused by increasing initiatives of local data providers and information intelligence.

Example of retail transaction search list with a predefined criteria. Additional data from the Polish commercial market are included in Annex 2.

Table 3: Retail transactions data from Polish regional cities

The screenshot shows the Comparables.pl website interface. On the left is a search filter panel with the following fields: Transaction type (Retail), Transaction subject (All), Country and Region (dropdowns), City and Street (text inputs), Transaction type (dropdown), Date from (III 2011) to (II 2012), Price in EUR from (10000000) to (50000000), Yield from (text input) to (text input), Rentable area from (text input) to (text input), Building name (text input), and Purchaser (dropdown). The main content area is titled 'Transactions matching search criteria' and shows 'Found 10 transactions. Showing 1 to 10'. Below this is a table with columns: Property type, Name, Location, Vendor, Purchaser, Date, and Price (EUR). The table lists four transactions, each with a small image of the property and a 'Retail' label below it.

Property type	Name	Location	Vendor	Purchaser	Date	Price (EUR)
Retail	Galeria Tęcza	Kalisz	Rank Progress	Blackstone Real Estate	I 2012	37 000 000
Retail	Komfort	Gorzów Wielkopolski	Wikana S.A.	Redevco	I 2012	11 500 000 Portfolio
Retail	CH Echo Pabianice	Pabianice	Echo Investment S.A.	PZU FIZ Sektora Nieruchomosci	IV 2011	22 887 230
Retail	CH Echo Bekchatów	Bekchatów	Echo Investment S.A.	Vousoka Ltd	II 2011	11 087 300

Source: Comparables.pl

---

## What is the yield?

Common interest for market players on developed, as well as on developing markets is the yield level. Within all markets, whether booming or in crisis, it always remains a relevant topic. It goes without saying that institutions from banks to strategic consultancies are monitoring yield levels within each market sector.

However, even if we allow that collating market level yields is possible, there is an issue in relation to determining what kind of yield is really reported.

Every analyst should possess basic skills regarding how to calculate the yields on the basis of rental income, estimated rental value and transaction price.

However, it remains a market practice even among market leaders to report transaction yields without further indication what kind of yield is reported. Obviously, the reason for this lies in the lack of data.

It is worth remembering that on developing markets it is very often the case that only initial yield (calculated on the basis of net operational income at the transaction date) is reported without further analysis. On the other hand, cash flow in post-transaction years usually remains unstable, which can be caused by lease termination, major renovations or step rents. Therefore, it is more appropriate to use equivalent yields in analysis. But the trick is, how to get all the data needed.

Another aspect of yield reporting is determining between gross and net yields, which might be confusing and misleading for foreign investors operating on some markets.

It is generally felt that research analysts should establish a common way of reporting and calculating standard.

Therefore, it is very often the case that an average price per square metre of rentable area remains the best benchmark.

Table 4: Commercial properties net prices per square meter of gross rentable area

1H 2013							
Prices in PLN	Poznan	Wrocław	Kraków	Łódź	Warszawa	Trójmiasto	Katowice
Office	9350	11950	9600	6200	12800	9250	7500
Retail	14000*	19000	-	13000*	15500	14000*	19950
Industrial	3225	-	3100*	2650*	2900*	2700*	2600*

Source: author's study based on recorded commercial transactions in 2013.

Obviously, net operating income is the determining factor of pricing on the commercial real estate market. However, when considering limited investment opportunities and similar demand profile of institutional investors', comparable data comprise a reliable benchmark.

### Spatial data

Last but not least, an interesting aspect of data gathering is its presentation.

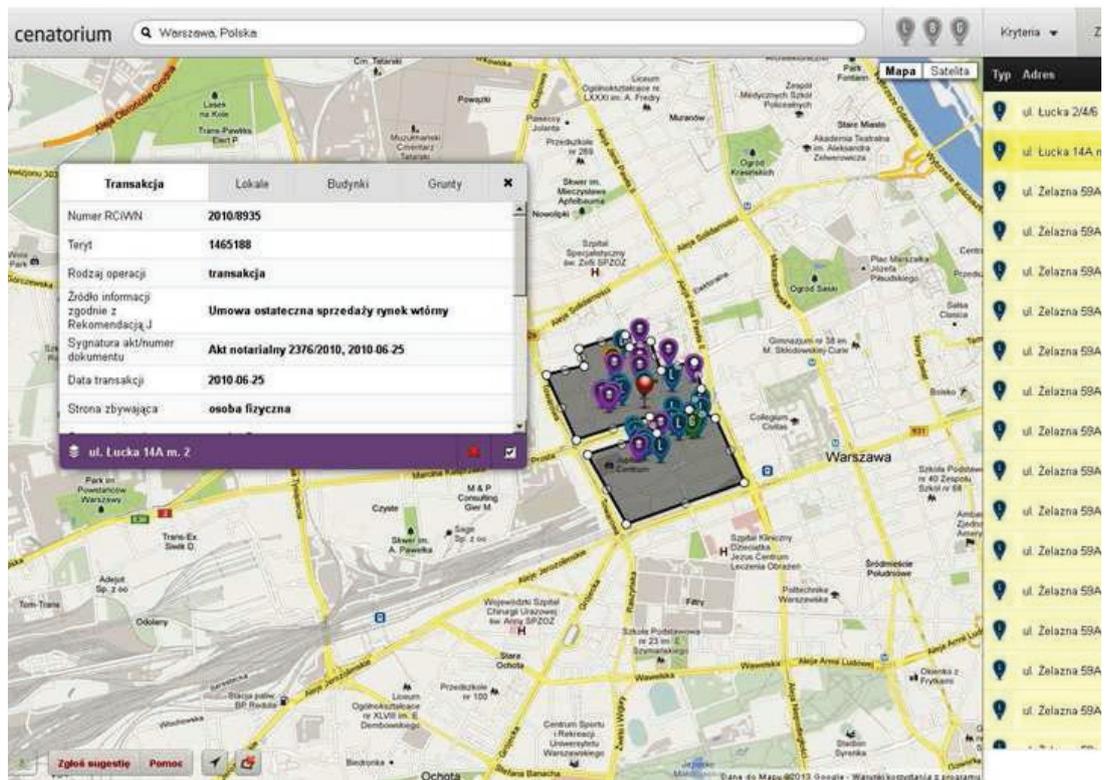
More particularly, it should be done by means of spatial GIS models, as exemplified in the model produced by the partnership of one of the market leaders in this field, CoStar with the GIS firm, ESRI (online: [www.giscale.com](http://www.giscale.com)). Such innovative modelling adds value to standard commercial property information.

CoStar Group provides useful data functionalities for agents, valuers and analysts, including vacancy rates, rental pace, void periods analysis, floorplans and comparables data, all of which is presented through GIS platforms. The initial steps in the market expansion of CoStar included field research photographers collecting data and photographing

properties, as well as street level video-imaging by 75 specially-equipped field research vehicles (online at <http://spatialnews.geocomm.com>). It is in addition to its by grid by grid aerial map imaging, of each of the researched properties, which currently covers almost 3 billion square meters of commercial properties in the United States and the United Kingdom.

The most recent example of the continuing evolution of the property related products, analysis and presentation is the use of flying drones, which have been successfully employed to enhance property imaging to an unprecedented scale.

Graph 2: Local property transactions presentation by one of the data providers



Source: Cenatorium.pl

One of the examples of local Polish developments in this field is the Cenatorium site, which combines property data collection and processing

with the use of data mining technologies. The site is currently targeted at banks and other financial institutions, but with the evolution of the product it is likely to attract wider clientele, including the consultancy sector of the real estate market. It is the first attempt in the CEE region to create a so-called "big data", an archive of information from different sources to be accessed quickly and efficiently.

### **In search of local indicators**

In market uncertainty, local consultants such as valuers have to resort to various forms of benchmarking their results.

The sales comparison approach (as one of the standardized valuation methods) remained the preferred valuation model. However, the residual approach was introduced widely in Poland at the turn of the century for the purposes of bank valuations. Although, it was widely criticised, especially by zealots of comparison method of valuation (referred to in Poland as "approach'), it became mandatory for many banking institutions when valuing development land.

However, uncritical approach to and reliance on comparable evidence may be very misleading. As Professor Ratcliffe put it "every comparable was a residual once".

Indisputably, it is a very sensible method, but when meticulously applied with market level assumptions, it remains less biased and uncertain than the comparable method.

Moreover, residual calculation combined with sensitivity analysis in the post-crisis reality is able to consider the uncertainty of certain value-affecting characteristics, and thus to quantify its impact on the market value.

One of the market indicators more widely used locally by valuers in Poland is land- to-capital value ratio. As it is described in *Modern*

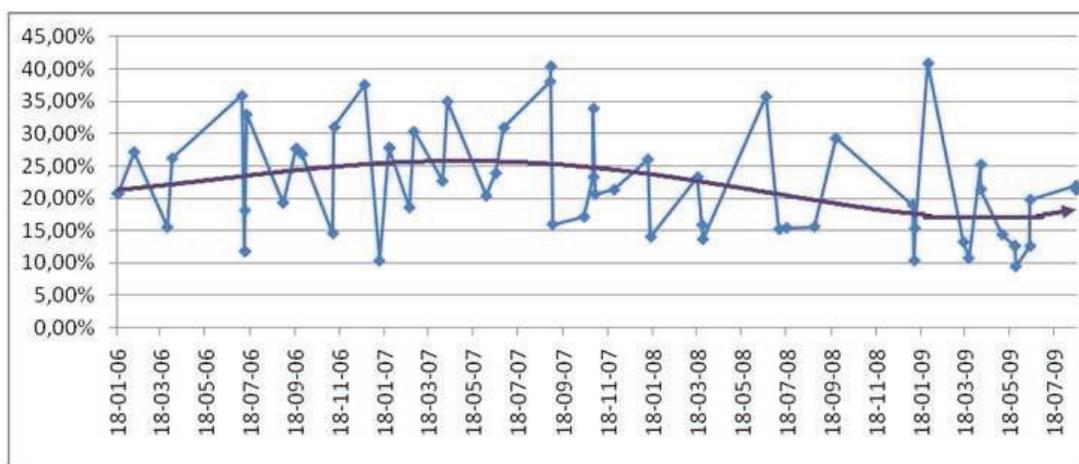
*Methods of Valuation*, in residual calculation a valuer is seeking to determine an overall size of the development “cake” which can then be “cut up” into the various slices needed for costs and profit to establish how much of the cake remains as the “land slice”.

Development land usually represents 10 to 25 per cent of total capital value of the completed development (Please see Annex 1), as in the case of industrial properties, as the land-to-capital value ratio constitutes approximately 10-15 per cent.

Interestingly, it appears that fluctuations of the above mentioned ratio in variable market conditions are statistically significant. Can it also serve as a “red flag” during property cycles?

When all other remaining sectors ratios (for which the arithmetical mean ranges from 20.75% to 23.13%), are depicted on one graph, some basic market rules may be noted.

Graph 3: Land/capital values ratio for combined sectors – timeseries



A visible drop in the ratio may be observed from the 2<sup>nd</sup> half of 2008. It followed a period of generally stable land-to-capital value ratios, which peaked in the III quarter of 2007. One might note that the ratio between land and capital value diminishes as the market deteriorates (please see Annex 1).

What might be of particular interest is the fact that the ratio appeared to go up again at the end of 2009. Probably, it can be treated as a herald of a market recovery of the Polish commercial market, which managed to avoid hitting a "second bottom".

It seems highly likely that, the land-to-capital value ratio shall remain on the agenda of valuers assessing both, vacant development land and projects under development.

Other markets are famous for using other market indicators. Calculating a percentage of reinstatement costs to market value might serve as one of the examples.

## 6. CONCLUSION

In summary, whether considering yields or price per gross leasable area, comparable market indicators are needed to provide market players with better market transparency, which is backed up by a more reliable market evidence.

Arguably, certain types of benchmarks, although requiring further research within wider samples and different data sources, can also serve as a “red flag” during property cycles.

The responsibility for developing market analysis of the commercial of the real estate sector carried by the private sector including the big consultancies and CoStar and Real Capital Analytics evolves and provides wider market perspectives. Although, market players are still waiting for coherent and universal methods of data analysis.

In the same time, modern analysis of commercial real estate has more to do with detective work than with mathematical analysis, which can remain relatively simple when collating the right data.

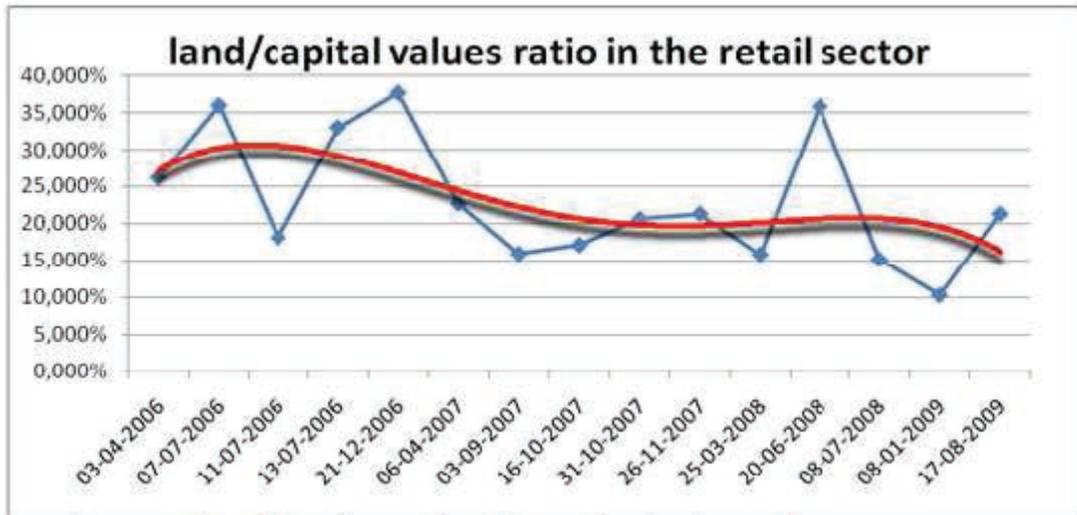
Therefore, lack of common reporting and data gathering standards continues to setback transparency of the commercial real estate markets worldwide.

## References

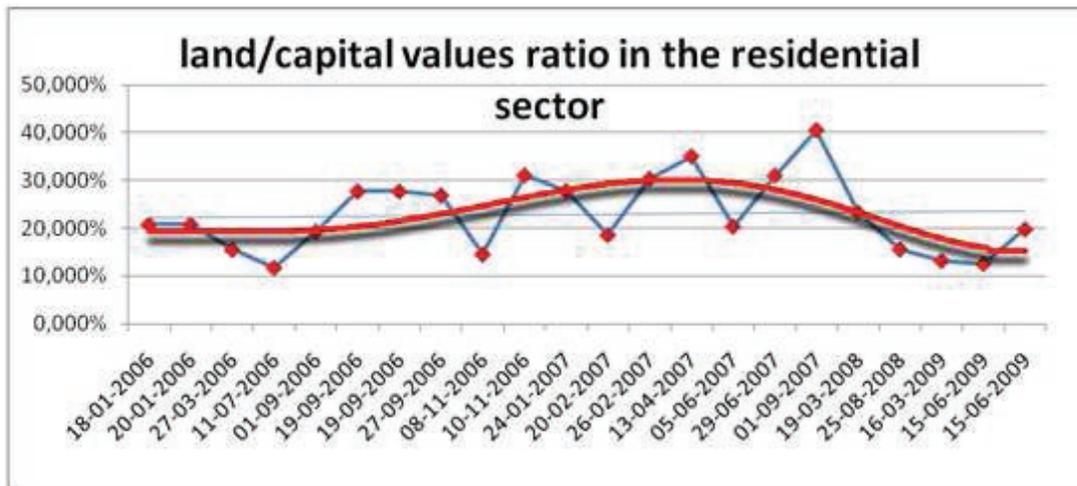
- CARR, Dennis (2003). *Mastering Real Estate Appraisal*. Chicago, Dearborn Real Estate Education.
- DOLIŃSKI, Wojciech *Research and Consultancy Project, The Assessment of Land Value*, Warsaw, 2009
- DOLIŃSKI, W. GRZESIK, Ch. *Real Estate Market Analysis – Global Outlook, XXI Annual Valuers Conference*, Miedzyzdroje, 2012
- GRZESIK, Ch. (2009) Valuers Simulate Softer Landing, *Real Estate Guide* 9 (1) ECE p. 51.
- RATCLIFFE, John et al (2004). *Urban Planning and Real Estate Development*. 2<sup>nd</sup> edition. London: Spon Press. Available at: <http://books.google.pl>. [Accessed 2 September 2013]
- SHAPIRO, Eric et al (2009). *Modern Methods of Valuation*. 10<sup>th</sup> edition. London: Estates Gazettes
- WHATLING, Arthur (2007). Valuing development land *The Journal RICS Commercial Property*, September/October 07.

## ANNEX 1

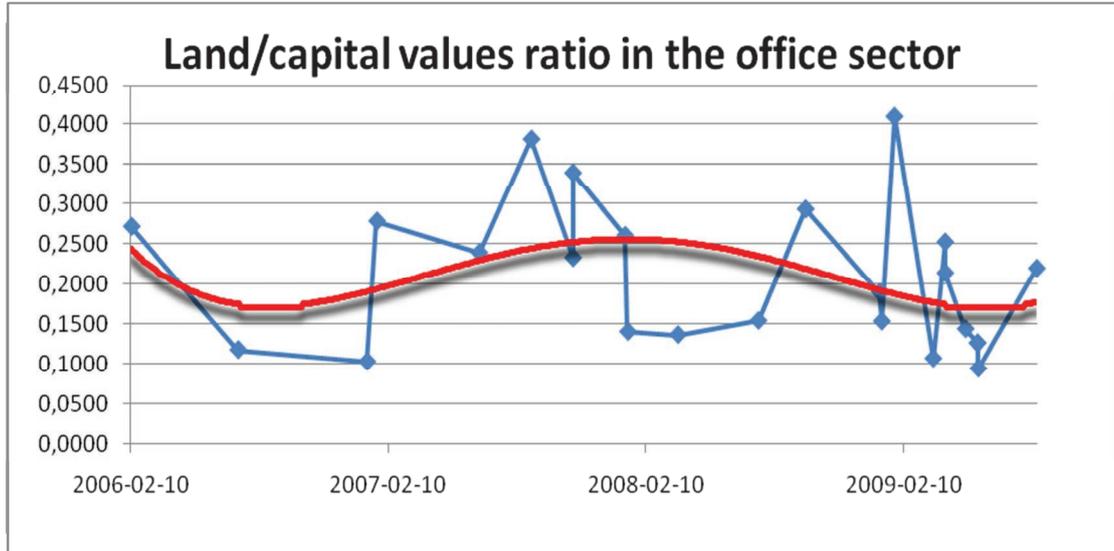
Research conducted in 2010 showed that ratio between land and capital value diminishes, as the market deteriorates (Dolinski, 2009). And this finding is confirmed by the available evidence from all four basic real estate market sectors.



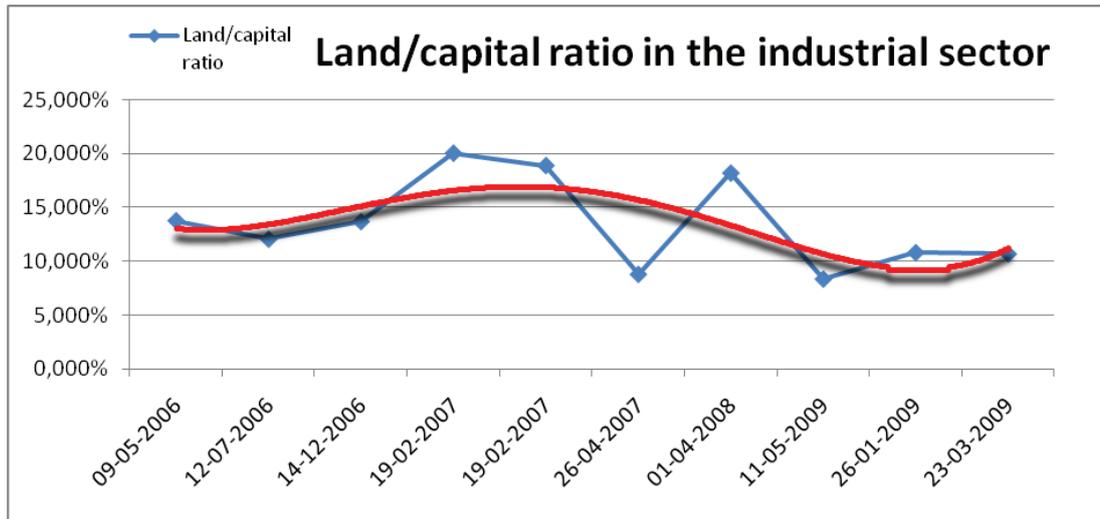
Graph 1: Land/capital values ratio – timeseries in the retail sector



Graph 2: Land/capital values ratio – timeseries in the residential sector



Graph 4: Land/capital values ratio – timeseries in the office sector

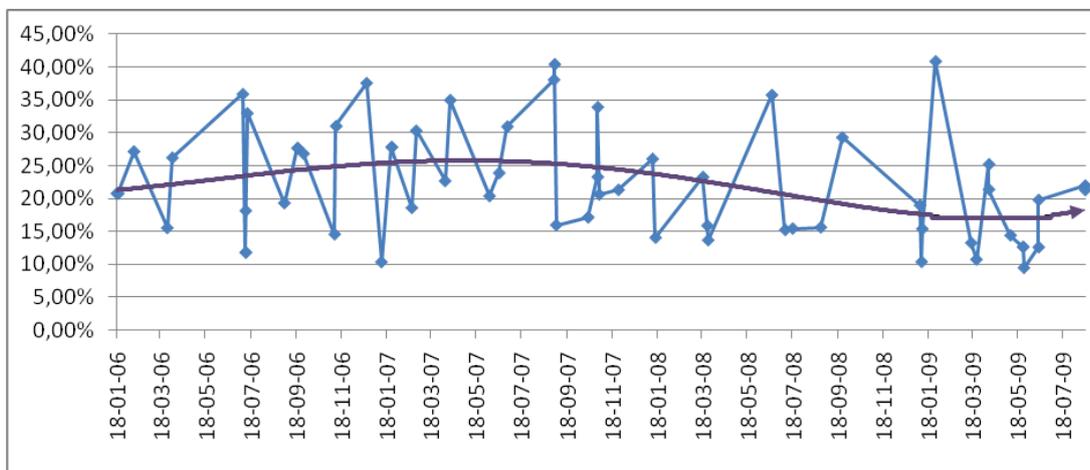


Graph 7: Land/capital values ratio – timeseries in the industrial sector

Obviously, within the warehousing sector, the market value of development land comprises a much lower share in the capital value of the completed development. Naturally, this is a consequence of built-up intensity as related to warehouse buildings. An additional consideration might also be in respect of market factors in the aspect of land scarcity. Industrial developments are usually constructed in areas where land does not have such unique location features as those found in the centre of the cities.

When all other remaining sectors ratios (for which the arithmetical mean ranges from 20.75% to 23.13%), are depicted on one graph, some basic market rules may be noted.

The results of the conducted research have led to astonishing conclusion that coefficient between land value and capital value oscillates around 0.85. In accordance with the Tables of probability, it clearly shows that the relation, is much stronger then it appears at first sight from the sample. Additionally, two suitable tests used for calculation of nonparametric correlation coefficients had been used in the Research, i.e. Spearman R and Kendall Tau.



Graph 8: Land/capital values ratio for combined sectors – timeseries

A visible drop in the ratio may be observed from the 2<sup>nd</sup> half of 2008. It followed a period of generally stable land-to-capital value ratios, which peaked in the III quarter of 2007. One might note that the ratio between land and capital value diminishes as the market deteriorates. It is worth mentioning that among many factors which influence the ratio between land and capital, the most important include supply and demand for both completed developments and construction service, as well as the availability of financing. Supply and Demand for completed developments dictate rent levels and apartment prices. Whereas, financing costs strongly influence the cost side of the residual equation.

## ANNEX 2

### Notable office deals in Poland dated 2013/2014

Property type ↑ ↓	Name ↑ ↓	Location ↑ ↓	Vendor ↑ ↓	Purchaser ↑ ↓	Date ↑ ↓	Price (EUR) ↑ ↓
 Office	Lipowy Office Park	Warszawa	CA Immobilien Anlagen AG	Kimberley	2014 I	108 000 000
 Office	Le Palais	Warszawa	Warimpex	IVG Warschau Fonds	2013 IV	31 000 000
 Office	Feniks	Warszawa	Europlan and Griffin Group	IVG Warschau Fonds	2013 IV	39 000 000
 Office	Atrium 1	Warszawa	Skanska Property Poland Sp. z o.o.	DEKA Immobilien GmbH	2013 IV	94 000 000
 Office	Intraco	Warszawa	PHN S.K.A.	PHN SPV 9 Sp. z o.o.	2013 IV	33 350 000

Source: Comparables.pl

An example of online database results depicting research within the industrial sector of transactions yielding between 8.40% and 9.25%.

**Transaction [?]** Property

Property type: Warehouse

Transaction subject: -----

Country: ----- Region: -----

City: [?] ----- Street: -----

Transaction type: -----

Date from (Q/Year): - 2011 to (Q/Year): II 2013

Price in EUR from: ----- to: -----

Yield from: 8,40 to: 9,25

Rentable area from: ----- to: -----

Building name: -----

Purchaser: -----

Vendor: -----

**Transactions matching search criteria**

Found 5 transactions. Showing 1 to 5

Property type ↑ ↓	Name ↑ ↓	Location ↑ ↓	Vendor ↑ ↓
 Warehouse	ProLogis Park Sosnowiec	Sosnowiec	ProLogis
 Warehouse	ProLogis Park Wrocław II	Wrocław	ProLogis European Properties (PEPR)
 Warehouse	Natolin 88 (GEFCO)	Grodzisk Mazowiecki	Invista European Real Estate Trust
 Warehouse	Panattoni Park Gliwice	Gliwice	Panattoni Europe

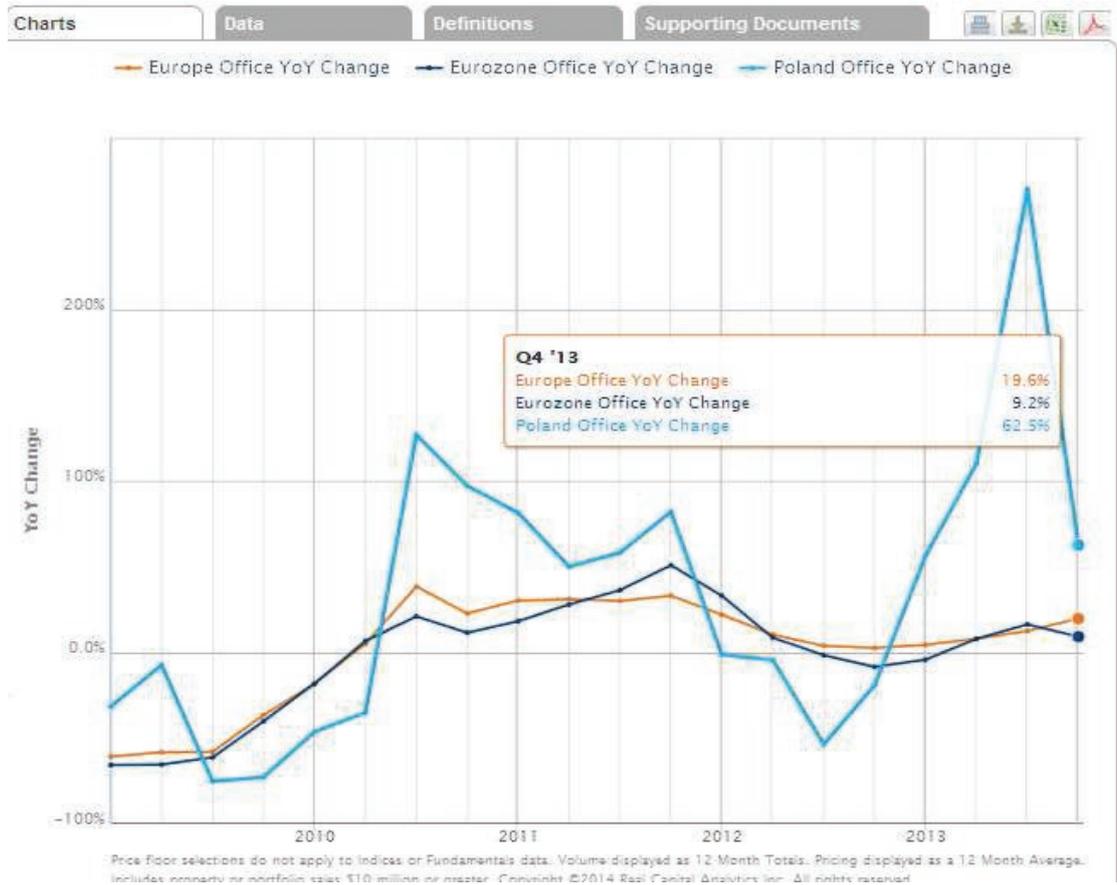
Source: Comparables.pl

### Top single property transactions in Europe closed in 2013

Property Name	Location	Price (€M)	m <sup>2</sup> /Units	PPm <sup>2</sup> /PPU	Type	Buyer
1 Citigroup Tower	London, GBR	€ 1,179.8	113,661	€ 10,380	Office	AGC Equity Partners
2 Chiswick Park	London, GBR	€ 939.5	176,510	€ 5,322	Office	China Investment Corp
3 Metropolis	Moscow, RUS	€ 859.5	205,000	€ 4,193	Retail	Morgan Stanley
4 Kings Cross Central Block A	London, GBR	€ 800.2	92,900	€ 8,614	Office	Google
5 Porta Nuova	Milan, ITA	n/a	290,000	n/a	Office	Qatar Investment Authority
6 White Square	Moscow, RUS	€ 742.1	74,000	€ 10,028	Office	O1 Properties
7 Shell Mex House	London, GBR	€ 715.1	50,998	€ 14,021	Office	Sirosa Ltd
8 Lotus City	Mamryi, RUS	€ 663.9	350,000	€ 1,897	Retail	God Nisanov JV Zarakh Iliev
9 White Gardens	Moscow, RUS	€ 590.1	63,300	€ 9,323	Office	Millhouse Capital
10 Ropemaker Place	London, GBR	€ 547.6	55,183	€ 9,923	Office	AXA Group JV SAFE
11 St Botolphs	London, GBR	€ 500.0	52,024	€ 9,611	Office	DekaBank
12 Devonshire House	London, GBR	€ 493.8	15,618	€ 31,616	Office	Ponte Gadea
13 Metropolis	Moscow, RUS	€ 463.01	205,000	€ 4,517	Retail	CalPERS JV Hines
14 Greenwich Peninsula Quays	London, GBR	n/a	607,028	n/a	Dev Site	Knight Dragon
15 Silesia City Center	Katowice, POL	€ 412.0	89,000	€ 4,629	Retail	Allianz JV Bank of China

Source: Real Capital Analytics

An example of RCA TrendTracker tool, enabling online monitoring of the investment activity within the office sector in Poland, benchmarking on European and Eurozone area dynamics.



Source: Real Capital Analytics

---

## 15. Measuring the effect of the real estate bubble: a house price index for Bilbao

M.J. Bárcena, P. Menéndez, M.B. Palacios, F. Tusell



M.J. Bárcena – Departament Statistics and Econometrics, University of the Basque Country (UPV/EHU), Bilbao (Spain).

P. Menéndez – University of New South Wales, Sydney (Australia).

M.B. Palacios – Departament of Statistics and O.R. Public University of Navarre (UPNA), Pamplona (Spain).

F. Tusell – Corresponding author.

Partial support from grants ECO2008-05622 (MCyT) and IT-347-10 (Basque Government) is gratefully acknowledged. We also acknowledge comments from participants in the NBP Workshop on Recent Trends in Real Estate Market and its Analysis.

---

## Article 15 Contents

1. Introduction .....	131
2. Housing prices and indices .....	132
3. Modelling price changes .....	138
4. Discussion .....	149
References .....	153

### Abstract

Houses are traded at relatively infrequent times and can hardly be standardized: two equally built and furnished houses may command widely different prices in the market on account of their different location or even orientation. Clearly, the computation of an index such as Laspeyres' cannot be contemplated, both because of lack of standardization and infrequent trading.

On the other hand, there is little market visibility; transaction prices are seldom published, further increasing the difficulties inherent to the construction of a house price index.

In this paper we use advertised (selling) prices as a proxy of transaction prices, and geographically weighted hedonic models to account for heterogeneity in quality and location. This gives a workable alternative to conventional approaches with the added benefit that it can provide price information in near-real time.

**Keywords:** house prices, spatial econometrics, semi-parametric models

**JEL Subject Classification:** C14, C81, R21

# 1 Introduction

Since the accession of Spain to the EEC in 1986, and in particular during the first years of the present century, housing prices have experienced substantial increases. This phenomenon has been observed in a large part of the western world, but in Spain has been exacerbated by a number of circumstances: monetary stability, with low or even negative real interest rates, easy borrowing, high economic growth and fiscal allowances for home buyers. Since the first years of this century it became commonplace to refer to the *real estate bubble*.

The financial crisis started in 2007 brought an abrupt end to this state of affairs. Credit tightened, house sales slowed down to almost a halt, and a number of building companies went bankrupt or had to file for protection under the Spanish bankruptcy law. When the slowdown propagated to all sectors and unemployment began to rise, more and more home owners have been unable to meet their mortgage payments. All this has led to substantial downward pressure in the residential property market, and price cuts have ensued.

As prices began to drop, widely different figures were given on how much they had already dropped (and by how much they would drop in the sequel). Most interestingly, widely different figures were given on the extent of the price increase while the bubble lasted.

Part of the discrepancy can be traced to the fact that different sources sometimes speak of different markets (it is clear, for instance, that second residences in coastal areas have been hit harder than urban properties in or near large cities). But even when speaking of roughly the same market, figures are so widely disagreeing so as to raise the question of how they were obtained and what they really measure.

This paper deals with the problem of measuring real estate property prices. We will use the housing market of the city of Bilbao as a test case, and discuss both results and methodological issues.

The remaining of this paper is organized as follows: Section 2 gives an overview of available sources of housing prices in Spain and the institutional framework. Section 3 introduces a semi-parametric and geographically weighted hedonic model which, among other things, produces a price index estimate. Section 4 compares our price index with related indices from alternative sources and draws some conclusions.

## 2 Housing prices and indices.

### 2.1 Methodological problems

The construction of price and quantity indices is one of the main tasks undertaken by all statistical offices. As such, it has been carefully scrutinized and the properties of different methods to construct indices are now well understood: see for instance Vogt and Barta (1996).

Laspeyres indices are one of the most popular choices. Given a basket made of quantities  $q_{i0}$  which at the base period  $s = 0$  had prices  $p_{i0}$ ,  $i = 1, \dots, I$ , the index for period  $s = t$  with base at period  $s = 0$  is computed as

$$I_0(t) = \frac{\sum_{i=1}^I p_{it} q_{i0}}{\sum_{i=1}^I p_{i0} q_{i0}} \quad (1)$$

This requires that items  $i = 1, \dots, I$  are well standardized and observable over time.

Houses are traded at relatively infrequent times and can hardly be standardized: two equally built and furnished houses may command widely different prices in the market on account of their different location or even orientation. Clearly, the computation of an index such as (1) is unfeasible, because no given set of houses are traded at regular intervals, and no exact replicates are available.

Both problems —lack of homogeneity of houses and irregular observation times— have been addressed in a variety of ways, which include the estimation of hedonic models, construction of “cells” of relatively similar houses and repeated sales methods (cf. Rodríguez López (2007), Widłak and Tomczyk (2010)).

**Hedonic models.** Hedonic models go back to at least the celebrated paper Court (1939). The paper Rosen (1974) is usually credited as laying down a sound theoretical basis. A useful collection of papers dealing with hedonic models in housing markets is Baranzini *et al.* (2008).

Hedonic models attempt to measure the contribution of individual characteristics to the total value of the house; each house is therefore regarded as a “basket” of characteristics like size, quality of the construction, age and the quality of the surroundings (e.g., clean air or noise). Location is typically included in the form of proxies which measure distance to transportation networks, municipal services, recreational areas, etc.

Since these proxies are not always observable and rarely capture fully all factors affecting the desirability of a given location, spatial effects are likely to remain; this can be introduced through the error structure, e.g.. Dubin (1988), as additional effects in the regression

(typically in the form of smooth surfaces over space, non-parametrically estimated, as in Hastie and Tibshirani (1991)) or by allowing the coefficients of the hedonic model to change in space: this approach, named geographically weighted regression (cf. Fotheringham *et al.* (2002)), will be followed in Section 3 (see also Kestens *et al.* (2005)).

Hedonic models have also been used to measure the effect of time on selling prices of homes, which is of particular interest here. The specification typically takes log price as the response variable in a regression model; the estimated coefficients of time dummies are used to measure the cumulative percentage of change in constant quality house prices up to and including the associated time period.

One advantage of hedonic models is that all home sales observed in a time period can be used in the estimation. By contrast, in repeated sales methods only houses which have been sold at least twice within the observed period can be used in the estimation. Quality adjustment is simple and automatic in repeated sales models<sup>1</sup>, as we consider differences in price of the same houses; the downside is that the sample size is reduced to houses with more than one trade.

## 2.2 Institutional framework

The market of new houses in Spain is split into the *free* and *protected* segments. House prices in the free segment are not subject to any restrictions whatsoever. The protected segment (*vivienda protegida*) is subject to ceilings in price, and is meant to guarantee a fraction of affordable houses for people who would otherwise be unable to buy one. A percentage of all new house starts must be in the protected segment<sup>2</sup>. After the first transmission, protected houses are traded much as free houses: the authorities have a preferential buying right of any protected house offered in the market, but only exceptionally has this right been exercised.

A feature of the real estate market is its opacity: prices really paid in real state transactions are rarely known with any degree of confidence. There are incentives to declare different (in general, lower) prices than those really paid, as taxes on the transactions are assessed on a per value basis<sup>3</sup>.

<sup>1</sup>Assuming house characteristics remain constant between consecutive sales dates.

<sup>2</sup>In the Basque Country, as of this writing, 20%.

<sup>3</sup>New houses are subject to the Value Added Tax (IVA), at a rate of 7% for most of the period analyzed, then raised to 8% of their value. In mid 2011 the IVA rate for new houses was lowered to 4%, in an attempt to boost the languishing housing market. Second hand houses pay a tax on transmissions of 6%. On the other hand, the seller must pay income tax on any increase of value of the sold property, although allowance is made for the effect of inflation. Both factors provide an incentive to understate the value of the transaction. Finally, there is a local tax on the increase of value of land, but this last is levied with no reference to the declared value of the transaction, and therefore provides no further incentive to distort declared prices.

Since real transaction prices are not readily observed, approximations are used. Declared prices, as mentioned, are usually different (and below) real prices. A different source of approximate prices are the value assessments made by specialized agents. These value assessments are collected nation-wide<sup>4</sup> and are one of the inputs for the Survey of Housing Prices<sup>5</sup> described below. For the free segment, these assessments of market value are usually required by banks from prospective buyers who wish to finance their buy with a mortgage<sup>6</sup>.

The use of market value assessments, however professionally made, raises several objections. First, there has been a tendency to overstate the value of properties in order to disguise loans which otherwise would appear insufficiently backed by the value of the mortgaged property. Second, it is not usually the case that a market value assessment is done on a house which is not mortgaged: this may lead to selection biases, as only part of the transactions are reflected in value assessments (cf. García Montalvo (2007)).

All these factors taken together explain that statistics on housing prices are sometimes difficult to reconcile, as they are based either on declared or assessed prices. We briefly describe some statistical sources of information next.

### 2.3 Statistical sources

There are several sources of statistical information on house prices in Spain. The Ministerio de la Vivienda, or Ministry of Housing (MV), and the Instituto Nacional de Estadística, or National Statistical Institute (INE), both provide information on housing market conditions and prices; the INE publishes the Índice de Precios de la Vivienda, or Housing Price Index (IPV), whose methodology is harmonized with that used in other countries of the European Union (EU). Other sources of information include real estate agencies and firms specializing in value assessment of real estate.

**Housing Prices Survey.** The Estadística de Precios de la Vivienda, or Housing Prices Survey (EPV), offers nation-wide statistical information on house prices and quantities provided by the MV<sup>7</sup>.

The MV addresses both the protected and free segments. The source data for the protected segment are the maximum legal prices. For the free segment, market value assessments are used. The assessed market value is defined as

---

<sup>4</sup>By ATASA, Asociación Profesional de Sociedades de Valoración.

<sup>5</sup>Estadística de Precios de la Vivienda, Ministerio de la Vivienda.

<sup>6</sup>The maximum amount of money to be lend is usually 80% of the assessed value of the property, although in the midst of the bubble 100% was sometimes granted.

<sup>7</sup>The information to follow is summarized from an undated report, Ministerio de la Vivienda (retrieved 10 June 2010).

“...the price at which a property could exchange hands at the date of the assessment, exclusive of taxes and commissions.”

Therefore, these are *not* prices at which transactions have taken place.

From the assessed market values, the MV produces indices using a bottom-up approach. Prices of homes of a given type (new or second hand, protected or free) are aggregated for each price stratum<sup>8</sup> and area and divided by the total built surface<sup>9</sup> to produce an average price in €/m<sup>2</sup>. These are further weighted and aggregated to produce indices for larger areas.

The price indices derived from the EPV are published quarterly.

**Assessment firms and real state agencies.** There are at least two price indices claiming fairly general coverage.

The Índice de Mercados Inmobiliarios Españoles, or Spanish Real Estate Market Index (IMIE) is an initiative of a private company<sup>10</sup> based on information from over 200.000 value assessments, collected by qualified staff with daily contact with the market. It is published monthly, and is a chained Laspeyres index, with base year 2001.

The spatial resolution tends to adapt to market segments rather than administrative boundaries: large cities, metropolitan areas, Mediterranean coast, Balearic and Canary Islands and Rest of municipalities. These market segments are seen as the core of the real state market, leaving into relative oblivion rural areas.

Aside from the use of assessed values rather than transaction prices, sample selection is to be feared, as assessment is more likely for properties to be mortgaged, and this may bias the sample towards some market and price segments.

Another private firm which publishes an index of the Spanish real state market is Sociedad de Tasación<sup>11</sup>. It produces an index of nominal prices per square meter, base year 1985. It is published twice per year and gives information for the whole of Spain as well as Madrid, Barcelona, Autonomous Communities and some large cities.

**Housing Prices Index.** The last addition to the sources of statistical information on housing prices is the IPV, an statistical operation whose aim and methodology is closest to our work. The IPV is a housing price index published quarterly by the INE with base year

<sup>8</sup>Seven strata are defined, with a top price of 1.050.000€; thus, the most expensive houses are excluded.

<sup>9</sup>Built surface (*superficie construida*) includes closed balconies, and structural vertical elements, as opposed to usable surface (*superficie útil*), which excludes those elements as well as 50% of balconies under roof; for a more precise description, see Ministerio de la Vivienda (retrieved 10 June 2010), p. 7. A useful rule of thumb is that usable surface is usually about 15% less than built surface.

<sup>10</sup>TINSA Consultancy and Environment, <http://www.tinsa.es>.

<sup>11</sup>It specializes in value assessments of all types, but particularly of real state assets. Present in the market since 1982, see <http://web.st-tasacion.es>.

2007. Its primary objective is to measure the evolution over time of house prices in both the new and second-hand free segments. It thus excludes prices in the protected segment. The IPV aims to serve as a comparable statistical source within the scope of the European Union harmonized statistics.

The document INE (2009) contains a fairly detailed discussion of the decisions made on primary sources to use. The basic information source retained is the database provided by the General Council of Spanish Notaries (Consejo General del Notariado). Virtually all house transactions are made with the intervention of a notary public. The transaction price along with the date, characteristics of the house (flat or single family house, built surface, new or second hand, parking space if available, etc.) is transmitted to the INE each month. Prices are therefore prices declared by both parties, buyer and seller, when signing the transfer of property. This need not be the real amount paid since, as explained in the previous section, there are incentives to understate (and, in some cases, to overstate) prices; however, it has been deemed the best alternative.

The IPV fully recognizes that quality of the houses transacted has to be taken into account. The approach followed involves the use of an hedonic model involving the variables mentioned above and additional information such as size of the municipality, type of environment, and the degree to which the location of the house is touristic. The dependent variable is the log price per square meter.

All house characteristics entering the hedonic model are qualitative<sup>12</sup>. The main effects they define along with selected interactions involve 157 parameters that have to be estimated each quarter. Each combination of explanatory variables defines one cell<sup>13</sup>. The model yields average price per cell which is then weighted into a general price index and price indices for different aggregations of cells<sup>14</sup>.

If we denote by  $\mathbf{x}_c$  the vector of values of the explanatory variables (main effects and interactions) defining cell  $c$ , and  $\ell_{i,c,q}$  the log price of house  $i$  in cell  $c$  at time (quarter)  $q$ , the model is:

$$\ell_{i,c,q} = \mathbf{x}_c \boldsymbol{\beta}^q + \epsilon_{i,c,q} \quad (2)$$

Stacking (2) for all observations corresponding to one quarter, we obtain:

$$\mathbf{L}_q = \mathbf{X}_q \boldsymbol{\beta}^q + \boldsymbol{\epsilon}_q \quad (3)$$

which is re-estimated every quarter  $q$ , yielding a time-varying vector of parameters  $\boldsymbol{\beta}^q$ .

<sup>12</sup>The built surface is categorized.

<sup>13</sup>For the years 2007 and 2008, nearly 52.000 such cells were used.

<sup>14</sup>The weights are given by the ratio between the value of houses transacted in the cell and the total value of houses transacted. Hence, selection biases may also creep in.

INE (2009) gives further details, concerning imputation of missing data, correction of heteroskedasticity and bias removal, so prices for each cell (rather than log prices) are approximately unbiased.

**Other sources of information.** The Basque Government publishes quarterly a bulletin, Departamento de Vivienda Obras Públicas y Transportes (2011), with a wealth of information (and commentaries) regarding trends not only in prices, but also on availability of housing, house starts, land prices, developable land, rents, etc.

Some web sites<sup>15</sup> which provide real estate advertising and services also publish average prices and price indices, as a by-product obtained from the information they have access to. Among them, [www.idealista.com](http://www.idealista.com), whose raw data we use in the models estimated in Section 3.

---

<sup>15</sup>See for instance <http://casas.facilisimo.com/preciometro> and <http://www.idealista.com/pagina/informes-precio-vivienda>.

### 3 Modelling price changes

#### 3.1 Data.

We have used data obtained from [idealista.com](http://idealista.com), a leading web site in the area of real state. Data consists of house offers in the city of Bilbao, from 2004-11-17 to 2011-05-04; in all, 7524, of which 6696 contained full street addresses that we have been able to geocode<sup>16</sup>. Data consists of a description of the property offered for sale, including selling price, location (often down to the street number, in other cases only street or district), square footage, floor, age of the building, number of bedrooms and bathrooms, parking space if available, availability of central or individual hot water and heating, elevator, etc. We emphasize that these are offered or selling prices, not transaction prices. While in principle transaction prices would be desirable, selling prices appear to be good proxies and have been used in a number of studies, including Kryvobokov and Wilhelmsson (2007), Henneberry (1998) and Pace *et al.* (2000).

Figure 1 shows a polygon outline with a map of Bilbao<sup>17</sup> as background, for reference. To avoid distraction, all subsequent plots of results will be made over the outline. Figure 2 shows the approximate location of districts referred to later.

We have fitted several models, whose relative merits are discussed next. It is remarkable that different specifications, using location information with various degrees of sophistication, still yield fairly similar estimates of the price index.

#### 3.2 Model 1.

Our first model is a standard hedonic model in which the response is  $\log(\text{Price}/\text{m}^2)$ . A categorical variable codes the district where each offered house is located; there are 13 such districts in Bilbao. Other (fixed) effects introduced are type of dwelling (flat, duplex, etc.), type of heating (or unavailability of heating, as the case may be), number of parking places, number of bathrooms, of bedrooms, whether there is elevator or not, the monthly cost of shared services and age of the building, coded in 7 categories.

The effect of time is modeled non-parametrically, as a cubic spline with 12 degrees of freedom. Thus, Model 1 can be written as:

$$\log(\text{Price}/\text{m}^2) = \sum_{i=1}^p \beta_i x_i + s(t) + \epsilon \quad (4)$$

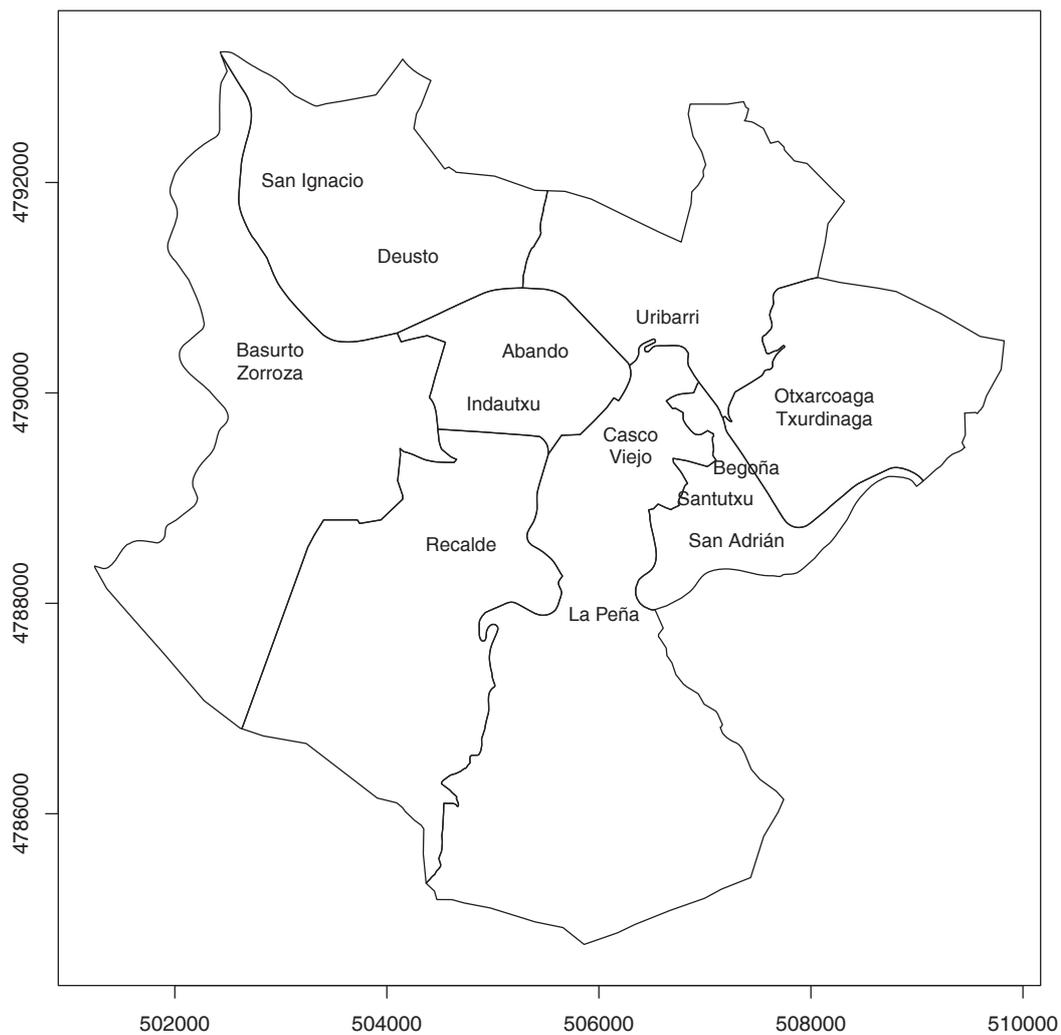
<sup>16</sup>But note that a smaller number of observations has been used to estimate the models below, as observations with missing variables may be discarded.

<sup>17</sup>From Google Maps, <http://maps.google.com>.

Figure 1: Outline of the city of Bilbao



Figure 2: Districts of Bilbao. The axis units are UTM coordinates within the sheet 30N, in meters.



where the  $x_i$  ( $i = 1, \dots, n$ ) are the observed regressors,  $s(t)$  is a cubic spline whose suitably normalized value provides our estimation of the price index and  $\epsilon$  is a random disturbance.

Models such as (4) can be readily fitted with off-the-shelf software. We have used the `gam` function in the R package `mgcv` (see Wood (2001) and Wood (2004)). Results are shown in Table 1. Since the response is measured in the log scale,  $\exp(s(t))$  enters as a factor in Price/m<sup>2</sup> and our estimation for the index with base 100 at time  $t_0$  is obtained<sup>18</sup> as

$$I(t) = 100 \times \frac{\exp(s(t))}{\exp(s(t_0))};$$

the profile of  $I(t)$  along with a 95% confidence interval are shown in Figure 3. We see that offered prices reached their maximum around the end of 2007, accumulating an increase of almost 25% since the beginning of 2005, then dropped until the first quarter of 2009 at which time they appear to have stabilized slightly above the early 2005 levels, only to later resume their drop.

Let  $P_i^*$  be the deflated price per square meter of house  $i$ . Next step is to fit a geographically weighted regression,

$$\log(P_i^*) = \sum_{j=1}^p \beta_{i,j} x_{ij} + \epsilon_i \quad (5)$$

where  $x_{ij}$  is the value of the  $j$ -th regressor for house  $i$ . Coefficients  $\beta_{i,j}$  are estimated for each regressor  $j$  at each house location  $i$ <sup>19</sup>.

The estimation is done by least squares, with observations weighted less as their distance to house  $i$  increases. Weighting is usually done with a two dimensional kernel, here an isotropic gaussian kernel whose bandwidth is selected by cross-validation. With the sample used, the bandwidth selected<sup>20</sup> (= standard deviation of the gaussian kernel) has been 350.8, which implies that observations 350.8 meters away are given weights about 6% of those right at the space point at which we estimate the coefficients (6%  $\approx$  0.0585, density of an isotropic bivariate normal density one standard deviation away from the mean). The cross-validated bandwidth is fairly small: in another setting, Can and Megbolugbe (1997) found that prices at one location are affected by prices within about a two mile radius.

Results of Model 2 are not presented here, as it is only a logical step towards Model 3.

<sup>18</sup>Note that both numerator and denominator in the fraction defining  $I(t)$  are biased. These biases are usually small in relative terms, and since they affect both numerator and denominator, their effect on  $I(t)$  is likely to be negligible; but they can be approximately corrected if desired.

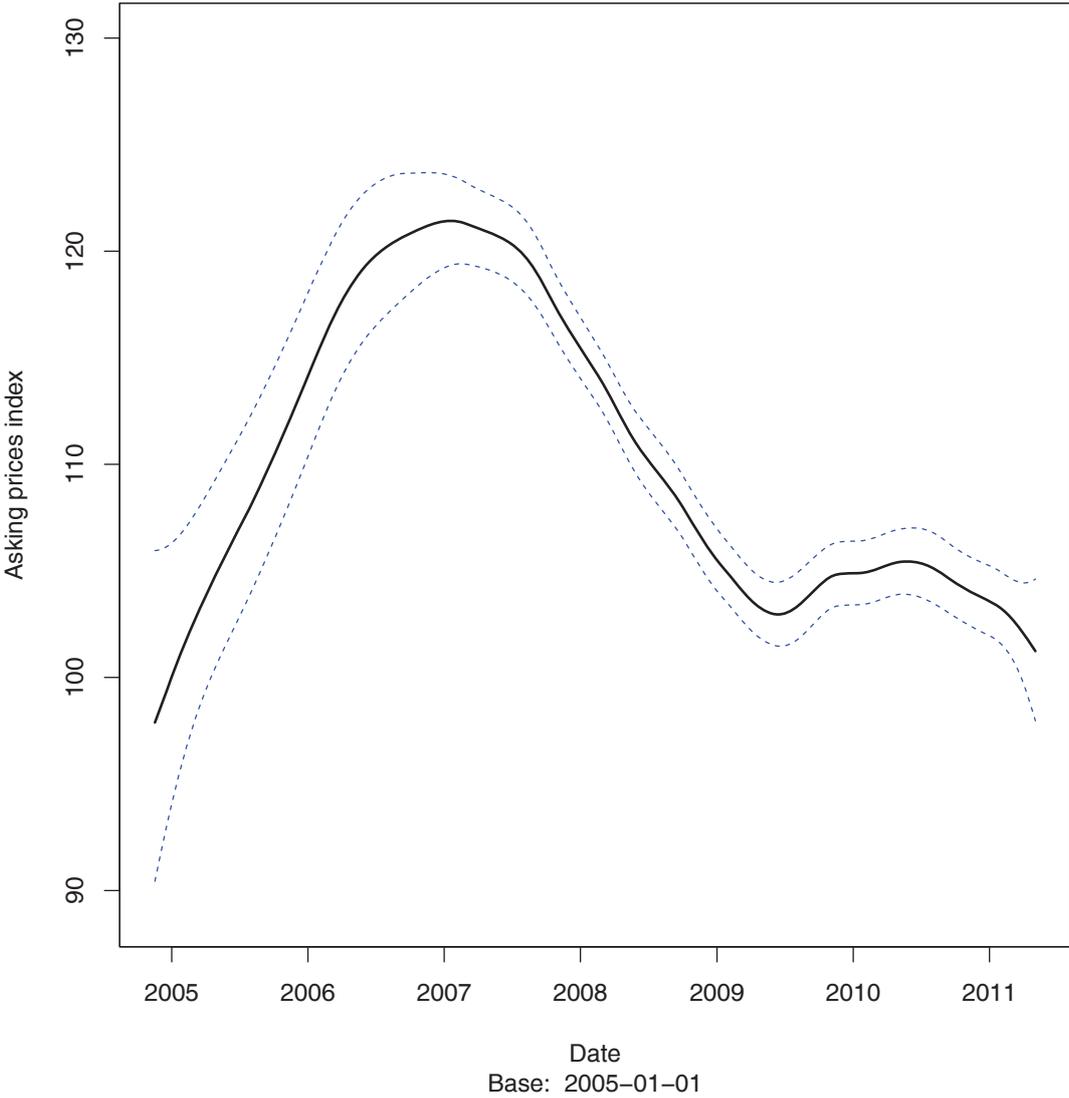
<sup>19</sup>Although for simplicity we are assuming that we estimate parameters only at locations where we observe a house, this need not be the case. We can estimate parameters at any point in space, irrespective of whether or not an observation is present there. In particular, we might estimate parameters over a grid of points to construct maps by smoothing or interpolation of the obtained values.

<sup>20</sup>As the geocoding is made to UTM coordinates in meters, the bandwidth is also in meters.

Table 1: Parameter estimates for Model 1.

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	9.6721	0.0503	192.34	0.0000
ExteriorTRUE	0.0120	0.0079	1.52	0.1290
Bedrooms	-0.0168	0.0040	-4.19	0.0000
Bathrooms	0.1055	0.0064	16.50	0.0000
Parking places	0.0737	0.0068	10.82	0.0000
log(Square footage)	-0.3384	0.0134	-25.18	0.0000
Community fees	0.0007	0.0001	7.32	0.0000
ElevatorTRUE	0.1859	0.0065	28.68	0.0000
<b>Type of heating:</b>				
colective	-0.1058	0.0132	-8.03	0.0000
unstated	0.0398	0.0073	5.48	0.0000
individual	-0.1413	0.0081	-17.48	0.0000
unavailable	-0.0012	0.0470	-0.03	0.9795
<b>Type of house:</b>				
duplex	-0.0505	0.0272	-1.86	0.0636
studio	-0.1764	0.0369	-4.78	0.0000
flat	-0.0411	0.0109	-3.77	0.0002
single house	-0.1565	0.0680	-2.30	0.0214
house in a block	-0.0141	0.0436	-0.32	0.7460
paired house	-0.0632	0.0564	-1.12	0.2620
<b>District:</b>				
basurto - zorroza	-0.0954	0.0091	-10.52	0.0000
begoña - santutxu	-0.0466	0.0100	-4.67	0.0000
casco viejo	0.0794	0.0124	6.41	0.0000
deusto	-0.0133	0.0114	-1.17	0.2439
ibaiondo	-0.1394	0.0088	-15.93	0.0000
indautxu	0.2196	0.0105	20.85	0.0000
otxarkoaga - txurdinaga	-0.1561	0.0156	-10.01	0.0000
rekalde	-0.1352	0.0153	-8.86	0.0000
san adrián - la peña	-0.0462	0.0085	-5.41	0.0000
san ignacio	-0.1573	0.0138	-11.37	0.0000
uribarri	-0.0724	0.0154	-4.70	0.0000
<b>Age of building:</b>				
< 5 years	0.0960	0.0120	8.02	0.0000
10-20 years	0.0292	0.0147	1.99	0.0469
20-30 years	-0.0112	0.0077	-1.46	0.1452
5-10 years	0.0685	0.0132	5.19	0.0000
unstated	0.0220	0.0078	2.82	0.0048
<b>Time trend:</b>				
s(x, 12)	-0.0001	0.0000	-20.83	0.0000

Figure 3: Housing price index for Bilbao: non-parametric estimate from Model 1. Base 100 January, 2005. Dotted lines give 95% point wise confidence intervals.



### 3.3 Model 3.

Model 2 in the preceding section is meant to capture the effect on prices of house characteristics in spatially dependent form at a given point in time. This is not our primary goal: what we want is to adjust prices for house quality and location, for the purpose of obtaining a constant quality price index. What we would like instead is a model such as

$$\log(P_{it}) = \sum_{j=1}^p \beta_{i,j} x_{ij} + s(t) + \epsilon_{it} \quad (6)$$

in which the linear coefficients  $\beta_{i,j}$  are as before hedonic coefficients spatially varying,  $P_{it}$  are current prices per square meter and  $s(t)$  is a smooth function capturing the evolution of prices. In a sense, (6) is a merger of Model 1 and Model 2, in that we have both a non-parametric estimate of the price level and spatially varying hedonic terms.

We are not aware of software enabling the fitting of models such as (6) off-the-shelf; however, given routines to estimate Models 1 and 2 it is fairly straightforward to implement a back-fitting estimation routine (cf. Hastie and Tibshirani (1991), § 4.4) for Model 3. The procedure can be sketched as follows:

---

STEP 0. Take as initial estimate  $s^{(0)}(t)$  of  $s(t)$  in (6) the smooth function estimated from (4) (or any other available approximation).

STEP 1. At iteration  $k$ , let  $\beta_{i,j}^{(k)}$  be the estimates of the  $\beta_{i,j}$  using geographically weighted regression to fit

$$\log(P_{it}) - s^{(k)}(t) = \sum_{j=1}^p \beta_{i,j}^{(k)} x_{ij} + \epsilon_{it}^{(k)} \quad (7)$$

STEP 2. At iteration  $k$ , obtain  $s^{(k+1)}(t)$  by smoothing over time the partial residuals  $\log(P_{it}) - \sum_{j=1}^p \beta_{i,j}^{(k)} x_{ij}$ .

STEP 3. For a pre-set tolerance  $\eta$ , if  $\max |s^{(k)}(t) - s^{(k+1)}(t)| < \eta$ , return as estimates the  $\beta_{i,j}^{(k+1)}$  and  $s^{(k+1)}(t)$  computed in the final iteration, otherwise return to Step 1.

---

Essentially, the back-fitting algorithm iterates Step 1 and 2, each time estimating the parametric or non-parametric part of the fit given the best current approximation of the other.

As it happens, convergence is attained in just three iterations; the starting (from (4), Model 1) and final (from (6), Model 3) estimates of the price index are shown in Figure 4. They are not very different from each other; the back-fitted index closely follows

its counterpart from Model 1 until mid 2006, and stays above ever since, albeit by a small amount. For the purpose of comparison, we plot also the IPV from INE (refer to section 2.3) for the whole of the Basque Country<sup>21</sup>. (Since the IPV is computed with base 2007 = 100, we have aligned their first published figure with the half sum of our two indices.)

Although the IPV is computed for a larger area and with different methodology, its profile is remarkably similar to our two indices. In particular, all of them nicely reflect a short-lived upturn in early 2010, likely due to the announced end of fiscal rebates for house buyers, which prompted a small surge of activity in the housing market.

The similar profiles of the IPV and asking prices indices are also significant in that they help to allay fears that they could measure entirely different things. One might hypothesize that asking prices are relatively sticky, and softening market conditions would surface rather in transaction prices (which are input to the IPV). This does not seem to be the case.

Model 3 produces a wealth of geographically disaggregated information. The parametric part (i.e., the right hand side of equation (7) evaluated with the estimated  $\beta_{i,j}^{(k)}$ ), suitably scaled, gives estimates  $\log(P_{it}/m^2) - s(t)$ , i.e. deflated log prices per square meter. Aligning values to the base period and transforming back to the original scale ( $\text{€}/\text{m}^2$ ), we get the values mapped in Figure 5. (We have made no attempt to correct biases due to the non-linear transformation.) The pattern in Figure 5 is clear; houses in the central part of the city command the highest prices, while peripheral districts of San Ignacio, Txurdinaga or Basurto exhibit lower prices.

It is also interesting to check the fit, which now is local. Since we compute a different regression for each location, we have local  $R^2$  coefficients that we have plotted in Figure 6. The fit of deflated log prices per square meter with the available regressors is modest, with the median  $R^2$  equal to 0.4494. There are localized areas of worse and better-than-average fit for which we have no obvious explanation.

In principle, local values for any of the  $\beta_{i,j}$  coefficients in equation (6) could be spatially represented. We could then have an estimate of how much e.g. an additional bedroom, existence of elevator or a garage, enhances the price of a house at each chosen location. This may be misleading: as pointed out in Wheeler and Tiefelsdorf (2005), local multicollinearity is usually a problem. Some remedies have been proposed, Wheeler (2006), which we have not attempted to implement, as for the purpose of price index construction we only require a good estimate of  $\sum_{j=1}^p \beta_{i,j}^{(k)} x_{ij}$  in (7), and are not concerned with apportioning that sum into the contributions of each predictor<sup>22</sup>.

<sup>21</sup>Obtained from their web page, <http://www.ine.es>, visited July, 14, 2011.

<sup>22</sup>The problem of collinearity in geographically weighted regression is further addressed in Bárcena *et al.*

Figure 4: Asking prices indices. The index from INE (base 2007) has been aligned with the average of our two indices at 2007-02-15, mid point of the first quarter for which it was published.

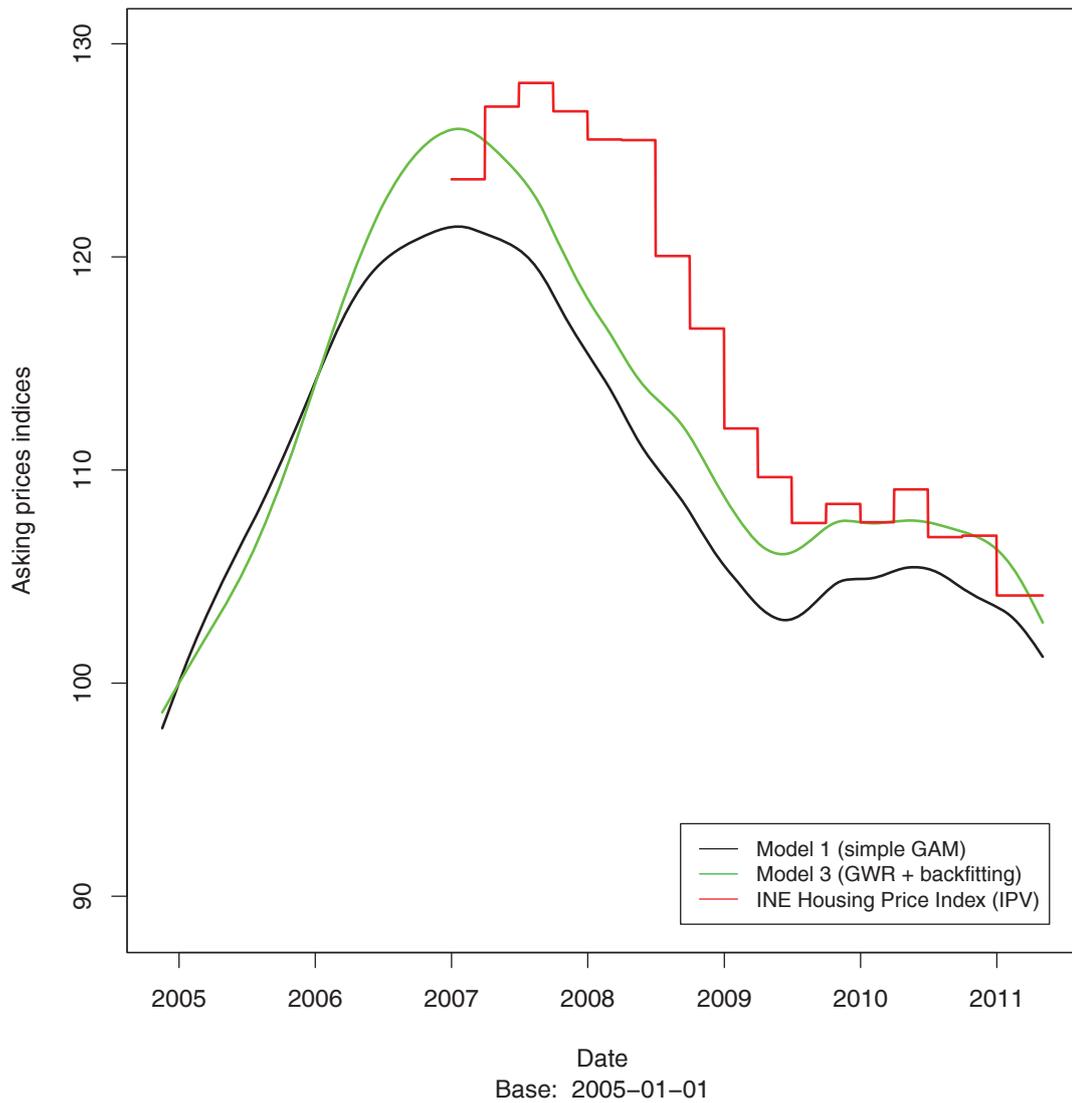


Figure 5: Fitted values in € per square meter at the base period (January, 2005).

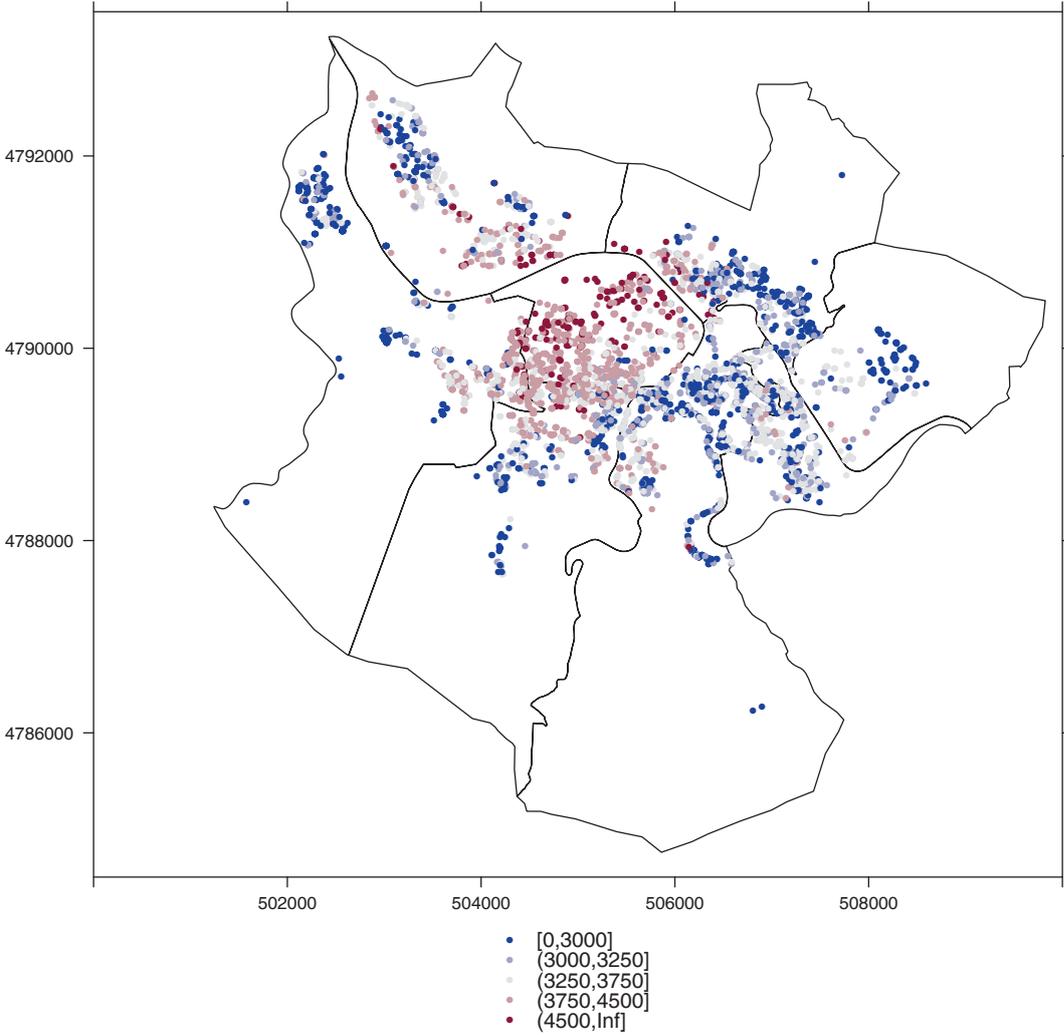
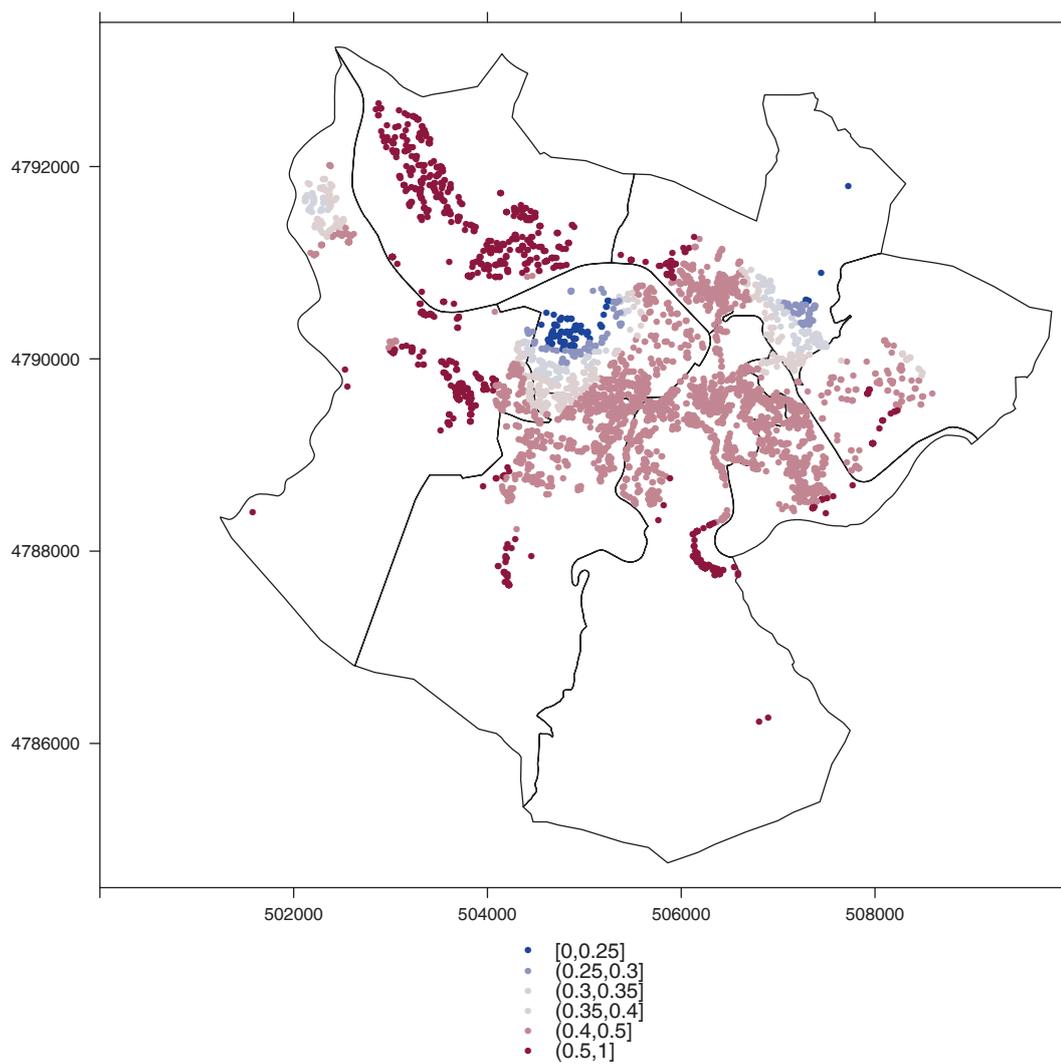


Figure 6: Local (partial)  $R^2$  values for the regression  $\log(P_{it}) - s^{(k)}(t) = \sum_{j=1}^p \beta_{i,j}^{(k)} x_{ij} + \epsilon_{it}^{(k)}$ .



## 4 Discussion

Our goal was to assess the feasibility of building a house price index with publicly available information on proxy variables of the true magnitudes to estimate. This is quite in keeping with the modern trend to exploit administrative records or automatically collected information to replace, to the extent possible, costly surveys.

The limited evidence shown in the precedent section is encouraging: with off-the-shelf software and limited human resources (basically spent in data cleaning and geocoding) we have been able to compute an index which captures remarkably well the patterns in the IPV<sup>23</sup>. The method is straightforward, conceptually simple to implement, understand and track. (Conceptual simplicity does not imply that the implementation is trivial; for an overview of some practical problems met see Bárcena *et al.* (2013).)

However, ours is but a possible approach to a much researched problem: for a collection of papers from a spatio-temporal vantage point see Biggeri and Ferrari (2010). Specifically related to housing prices and close to our approach are, among many others, Borst (2008), Geniaux and Napoléone (2008), Clapp (2004), Meese and Wallace (1997) Bourassa *et al.* (2006), McMillen (2011), Brunauer *et al.* (2010), Brunauer *et al.* (2012) and McMillen and Redfearn (2007). Comments on some of their respective approaches follow, to highlight similarities or differences with the approach followed in Model 3; Table 2 gives a summary. We can make no attempt to even list all contributions in this area, which has produced a copious literature.

McMillen (2011) goal is to examine the effect over time on house prices of the distance to “employment sub-centers” in the city of Chicago. He uses a repeated sales estimator and a smooth function to capture the effect of time. His model takes the form

$$Y_{i,t} - Y_{i,s} = g(t) - g(s) + u_{i,t} - u_{i,s}, \quad (8)$$

where  $Y_{i,t}$  is the (log) price of house  $i$  at time  $t$ ,  $u_{it}$  the corresponding random term and  $g(t)$  is a smooth function of time given by

$$g(t) = \alpha_0 + \alpha_1 z + \alpha_2 z^2 + \sum_q (\lambda_q \sin(qz) + \gamma_q \cos(qz)), \quad (9)$$

with  $z = 2\pi t / \max(t)$ . As compared with our model, he does not need to fit a hedonic

---

(submitted for publication)

<sup>23</sup>It is important not to be over-optimistic when interpreting Figure 4. One might be tempted to claim that our index appears to anticipate movements in the IPV. This would not be a fair claim, as the IPV is computed with only past and present information, while our index uses also *future* information in the smooth.

model, as the effect of the quality of house  $i$  cancels in the difference  $Y_{i,t} - Y_{i,s}$  (assuming the characteristics of the house  $i$  remained constant from time  $s$  to time  $t$ ); the downside is that only houses with repeated sales in the sample can be used (with a possible selection effect, as poor quality houses might be more likely to enter the market repeatedly). To capture the effect of time, a second order polynomial plus a combination of periodic functions is used. The profile of  $g(t)$  suitable normalized would give an estimate of a price index similar to ours, although for a given number of degrees of freedom we regard the penalized spline approach we have used as more flexible.

Geniaux and Napoléone (2008) are not concerned with the effect of time, but rather spatial effects and in particular the distance effect to urban center. In their paper they examine the effect of distance to Avignon on prices of rural properties in the south of France. They consider a variety of models, like

$$Y_i = \beta X_i + Z_i * s(u_i, v_i) + s(u_i, v_i) + \epsilon_i \quad (10)$$

where  $Y_i$  is again the log price of house  $i$ ,  $X_i$  captures characteristics that have a fixed (across space) effect, and  $Z_i$  is a variable whose effect is space-dependent; it is multiplied by a “smooth” function of geographical coordinates to capture such dependency. They go on considering alternatives such as generalized additive models (GAM), geographically weighted regression (GWR) and mixed GWR (MGWR). They conclude that

“GAM fits better than MGWR, is even more flexible in articulating stationary and non stationary coefficients, works well with a big sample and makes investigating non linearity easy.” (p. 125)

While they are not concerned with the problem of modelling the effect of time (which enters some of their models only in the form of year dummies), their models are close to ours; in their Section 5.4.3 they propose an estimation method for the MGWR model which is also close to the back-fitting algorithm sketched above for Model 3.

Close to this approach is also Clapp (2004), which proposes the model,

$$Y_{it} = \beta X_i + s(u_i, v_i, t) + \epsilon_{it} \quad (11)$$

where, as before,  $Y_{it}$  is the log of the transaction price,  $X_i$  is a vector of attributes of the  $i$ -th house in the sample,  $\beta$  the vector of implicit hedonic prices, and  $s(\cdot)$  a function of latitude, longitude *and* time. They consider fixed betas in the hedonic part and a smooth function in time and space (estimated by local polynomial regression, using a product kernel) to account for trends in time and space. The estimated  $s(u_i, v_i, t)$  as a function of time for given  $u_i, v_i$  gives a *local* price index. The estimation of (11) is made in a two-step

Table 2: Summary and comparison of some models and indices discussed

MODEL OR INDEX	QUALITY ADJUSTMENT	SPATIAL EFFECTS	TIME EFFECTS
Geniaux and Napoleone(2008): $Y_i = \beta X_i + Z_i s(u_i, v_i) + s(u_i, v_i) + \epsilon_i$	hedonic, fixed $\beta$	smooth $s(u_i, v_i)$	None
McMillen(2011): $Y_{i,t} - Y_{i,s} = g(t) - g(s) + \epsilon_{i,t} - \epsilon_{i,s}$	repeated sales	not explicitly	$g(t)$
Clapp(2004): $Y_{it} = \beta X_i + s(u_i, v_i, t) + \epsilon_{it}$	hedonic, fixed $\beta$	joint smooth $s(u_i, v_i, t)$	
Meese and Wallace(1997): $Y_{i(t)} - \bar{Y}_t = G(x_{i(t)}) + \epsilon_{i(t)}$	hedonic	LWR <sup>†</sup> in attributes	estimated for each $t$
Housing Prices Survey (IPV): $Y_{i,c,t} = X_c \beta_t + e_{i,c,t}$	hedonic, variable $\beta$	not explicitly	estimated for each $t$
Model 3 (Section 3.4 above): $Y_{it} = \beta_i X_{it} + s(t) + \epsilon_{it}$	hedonic, variable $\beta$	GWR <sup>‡</sup>	smooth $s(t)$

† LWR = Locally weighted regression.

‡ GWR = Geographically weighted regression.

algorithm, using theory developed by Robinson (1988) and Stock (1989).

Meese and Wallace (1997) compare several methods, providing a wealth of insight on their respective merits. They propose yet another model,

$$Y_{i(t)} - \bar{Y}_t = G(x_{i(t)}) + u_{i(t)}; \quad (12)$$

Equation (12) is estimated for each quarter;  $Y_{i(t)}$  is the log price of the  $i$  house traded at quarter  $t$ , and  $\bar{Y}_t$  is the average log prices of all houses traded in that same quarter.  $G(x_{i(t)})$  is the hedonic part, estimated by locally weighted regression (LWR); weighting is done not in terms of geographical distance, but rather distance in the attribute space to the attributes median. The implicit prices of the attributes are then used to correct  $\bar{Y}_t$  with the valuation of the attributes of the current or first quarter. This gives Laspeyres and Paasche indices whose geometric mean produces a final Fisher's ideal index.

McMillen and Redfearn (2007) contains an enlightening discussion on kernel regression, conditionally parametric regression (CPAR) and locally weighted regression (LWR), which they show encompasses the previous methods and GWR as particular cases. An interesting

approach that uses *both* sales prices and appraisal prices is described in Bourassa *et al.* (2006), which also contains a comparison among several versions of hedonic models, repeated sales methods and their SPAR (Sale Price Appraisal Ratio) method.

Brunauer *et al.* (2010) is concerned with rent levels rather than property prices. It develops a model with “spatial scaling factors”: the effect on rents of non-categorical variables such as square footage or year of construction is captured by way of smooth functions multiplied by scaling factors, which are district-dependent, and constrained to vary smoothly by penalizing differences among neighbouring districts. The time dimension is not involved. Brunauer *et al.* (2012), on the other hand, develops a price index similar to ours in the use of a non-parametric smooth function to capture the price evolution in time. District dummies are used, with random effects modeling unexplained spatial heterogeneity.

Our Model 3 produces an index for the whole area sampled, rather than indices for each particular location —like e.g. Clapp (2004). Where other models, such as Geniaux and Napoléone (2008), use smooth nonparametric functions in geographical coordinates to account for spatial effects, our model accounts for spatial variation through GWR of hedonic coefficients, with the  $s(t)$  term smoothing along the time dimension, much like McMillen (2011). Our model is geared towards production of an index for a region, rather than the discovery of sub-regions with different price trends.

Aside from the papers mentioned, there is a huge literature on spatio-temporal processes, potentially useful for our purposes. An up-to-date comprehensive overview is given in Cressie and Wikle (2011); see also Banerjee *et al.* (2003) and Gelfand *et al.* (2010). Specific applications to housing prices already exist that make use of this approach, e.g. Gelfand *et al.* (2004). For the purpose of computing price indices, we find hierarchical models a highly attractive alternative, for the reasons outlined in Cressie and Wikle (2011), § 2.1; they afford a nice separation between the “process model”, describing how relevant magnitudes (or state variables) evolve, and the “data model”, describing how observations are generated. Since our concern is to estimate (unobservable) latent variables such as the price level, this separation is meaningful; we may confine the latent variables to the “process model” and build the data model on top of it. Although not directly concerned with price indices, Brunauer *et al.* (2013) offers an example of such an approach.

Specialized software exists that can be used to obtain posterior distributions of the magnitudes of interest through MCMC, if one is willing to follow a Bayesian approach (e.g. Smith *et al.* (2008)) although this is computationally demanding for the problem sizes we envision. Alternative methods like integrated nested Laplace approximations (cf. Rue *et al.* (2009)) might be used instead.

## References

- S. Banerjee, B.P. Carlin, and A.E. Gelfand. *Hierarchical Modeling and Analysis for Spatial Data*. Chapman & Hall/CRC, 2003.
- A. Baranzini, J. Ramirez, C. Schaerer, and P. Thalmann, editors. *Hedonic Methods in Housing Markets*. Springer Verlag, New York, NY, 2008.
- M.J. Bárcena, P. Menéndez, M.B. Palacios, and F. Tusell. A real-time property value index based on web data. In Yanchang Zhao and Yonghua Cen, editors, *Data Mining Applications in R*, pages 273–298. Academic Press, 2013.
- M.J. Bárcena, P. Menéndez, M.B. Palacios, and F. Tusell. Alleviating the effect of collinearity in geographically weighted regression. submitted for publication.
- L. Biggeri and G. Ferrari, editors. *Price Indexes in Time and Space*. Springer Verlag, 2010.
- R.A. Borst. Time-Varying Model Parameters: Obtaining Time Trends in a Hedonic Model Without Specifying Their Functional Form. *Journal of Property Tax Assessment & Administration*, 6(4):29–36, 2008.
- S Bourassa, M Hoesli, and J Sun. A simple alternative house price index method. *Journal of Housing Economics*, 15(1):80–97, March 2006.
- W.A. Brunauer, S. Lang, P. Wechselberger, and S. Bienert. Additive hedonic regression models with spatial scaling factors: An application for rents in Vienna. *Journal of Real State Finance Economics*, 41:390–411, 2010.
- W.A. Brunauer, W. Feilmayr, and K. Wagner. A new residential property price index for Austria. *Statistiken*, Q3:90–102, 2012.
- W.A. Brunauer, S. Lang, and N. Umlauf. Modelling house prices using multilevel structured additive regression. *Statistical Modelling*, 13(2):95–123, 2013.
- A. Can and I. Megbolugbe. Spatial Dependence and House Price Index Construction. *Journal of Real Estate Finance and Economics*, 222:203–222, 1997.
- J.M. Clapp. A Semiparametric Method for Estimating Local House Price Indices. *Real Estate Economics*, 32(1):127–160, March 2004.
- A.T. Court. *Hedonic Price Indexes with Automotive Examples*, pages 99–117. General Motors Corporation, 1939.

- N. Cressie and C.K. Wikle. *Statistics for Spatio-Temporal Data*. Wiley, 2011.
- Departamento de Vivienda Obras Públicas y Transportes. Eskaintza inmobiliarioa. Oferta inmobiliaria, 2011.
- R.A. Dubin. Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms. *The Review of Economics and Statistics*, 70(3):466–474, 1988.
- S. Fotheringham, M. Charlton, and C. Brunson. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Wiley, 2002.
- J. García Montalvo. ¿Cuánto aumenta realmente el precio de la vivienda? *Índice. Revista de Estadística y Sociedad*, 22:10–11, 2007.
- A.E. Gelfand, M.D. Ecker, J.R. Knight, and C.F. Sirmans. The Dynamics of Location in Home Price. *The Journal of Real Estate Finance and Economics*, 29(2):149–166, September 2004.
- A.E. Gelfand, P.J. Diggle, Montserrat Fuentes, and Peter Guttorp, editors. *Handbook of Spatial Statistics*. CRC Press, 2010.
- G. Geniaux and C. Napoléone. Semi-parametric tools for spatial hedonic models: An introduction to mixed geographically weighted regression and geadditive models. In *Hedonic Methods in Housing Markets*, pages 101–127. Springer Verlag, 2008.
- T.J. Hastie and R.J. Tibshirani. *Generalized Additive Models*. Chapman & Hall, London, 2nd. edition, 1991.
- J. Henneberry. Transport Investment and House Prices. *Journal of Property, Valuation and Investment*, 16(2):144–157, 1998.
- INE. Índice de Precios de Vivienda. Metodología. Technical report, Instituto Nacional de Estadística (INE), 2009.
- Y. Kestens, M. Thériault, and F. Des Rosiers. Heterogeneity in hedonic modelling of house prices: looking at buyers' household profiles. *Journal of Geographical Systems*, 8(1):61–96, November 2005.
- M. Kryvobokov and M. Wilhelmsson. Analysing location attributes with a hedonic model for apartment prices in Donetsk, Ukraine. *International Journal of Strategic Property Management*, pages 157–178, 2007.
- J. LeSage and R. Pace, editors. *Advances in Econometrics*. Emerald, New York, NY, 2011.

- Ministerio de la Vivienda. Notas metodológicas. Undated report, retrieved 10-June-2010.
- D.P. McMillen and C. Redfearn. Estimation, Interpretation, and Hypothesis Testing for Nonparametric Hedonic House Price Functions. 2007.
- D.P. McMillen. *Employment subcenters and home price appreciation rates in metropolitan Chicago*, pages 237–257. In LeSage and Pace LeSage and Pace (2011), 2011.
- R.A. Meese and N.E. Wallace. The Construction of Residential Housing Price Indices: A Comparison of Repeat-Sales, Hedonic-Regression, and Hybrid Approaches. *Journal of Real Estate Finance and Economics*, 73:51–73, 1997.
- R.K. Pace, R. Barry, O.W. Gilley, and C.F. Sirmans. A method for spatial-temporal forecasting with an application to real state prices. *International Journal of Forecasting*, 16:229–246, 2000.
- P.M. Robinson. Root-n-consistent semiparametric regression. *Econometrica*, 56:931–954, 1988.
- J. Rodríguez López. Los índices de precios de la vivienda. Problemática. *Indice. Revista de Estadística y Sociedad.*, 22:14–16, 2007.
- S. Rosen. Hedonic prices and implicit markets: Product differentiation in pure competition. *The Journal of Political Economy*, 82:34–55, 1974.
- H. Rue, S. Martino, and N. Chopin. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392, April 2009.
- B.J. Smith, Jun Yan, and M.K. Cowles. Unified geostatistical modeling for data fusion and spatial heteroskedasticity with R package ramps. *Journal of Statistical Software*, 25(10):1–21, 2008.
- J.H. Stock. Nonparametric policy analysis. *Journal of the American Statistical Association*, 84(406):567–575, 1989.
- A. Vogt and J. Barta. *The Making of Tests for Index Numbers*. Physica-Verlag, 1996.
- D. Wheeler and M. Tiefelsdorf. Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems*, 7(2):161–187, June 2005.

- D.C. Wheeler. *Diagnostic tools and remedial methods for collinearity in linear regression models with spatially varying coefficients*. PhD thesis, Ohio State University, 2006.
- M. Wiłlak and E. Tomczyk. Measuring price dynamics: evidence from the Warsaw housing market. *Journal of European Real Estate Research*, 3(3):203–227, 2010.
- S.N. Wood. mgcv: GAMs and generalized ridge regression for R. *R News*, 1(2):20–25, June 2001.
- S.N. Wood. Stable and efficient multiple smoothing parameter estimation for generalized additive models. *Journal of the American Statistical Association*, 99:673–686, 2004.
- Yanchang Zhao and Yonghua Cen, editors. *Data Mining Applications in R*. Academic Press, 2013.

---

## 16. Recent developments in house prices in Macedonia

**Branimir Jovanovic**



Branimir Jovanovic – National Bank of the Republic of Macedonia.

The views expressed in this paper are those of the author and do not have to represent the views of the National Bank of the Republic of Macedonia.

---

## Article 16 Contents

I. Abstract .....	160
II. Introduction .....	161
III. Hedonic house price index for Macedonia .....	162
IV. General movements in the index .....	164
V. Econometric analysis of Macedonian house prices .....	165
V.A. The model .....	165
V.B. Variables and data .....	165
V.C. Econometric analysis .....	167
V.D. Explaining recent movements in house prices .....	169
V.E. Comparison between actual and equilibrium price .....	170
VI. Conclusion .....	171
VII. References .....	172

## I. ABSTRACT

We overview developments in house prices in Macedonia in the period 2000-2013, with a special focus on the movements during the Great Recession. House prices in Macedonia exhibit typical cyclical behaviour, where periods of growth are followed by periods of stagnation/decline. Price developments are well explained by standard fundamentals, such as income, mortgages and supply of new housing, and during these 14 years houses have not shown signs of substantial and persistent mispricing. Real house prices have declined by approximately 17 percent since the onset of the global financial crisis (end of 2008). The decline is in accordance with the lower demand during this period (due to the stagnation in real income and the slowdown in the credit growth) and the higher supply of new housing.

**JEL classification:** R31

**Keywords:** housing market, house prices, Macedonia, financial crisis, Great Recession

## II. INTRODUCTION

The housing market is very important for the financial stability, and the economic activity in general, as evidenced by the notion that the recent global financial crisis originated in the US housing market. Hence, central banks should actively monitor developments in the housing market, in order to detect potential bubbles (i.e. price increases not justified by fundamentals, but only by expectations of future increases in prices; see Stiglitz, 1990) and to take appropriate respective measures.

In this paper we will present an overview of the developments in the house prices in Macedonia, from the perspective of the National Bank of the Republic of Macedonia (NBRM). First, we will describe the hedonic house price index, constructed by the NBRM. Then we will briefly overview general trends in the index. Next, we will analyze the index using econometric techniques, in order to see what has been driving house prices. Finally, we will discuss movements in the house prices after the global financial crisis and will compare the actual house price with the equilibrium one.

### III. HEDONIC HOUSE PRICE INDEX FOR MACEDONIA

When the subprime crisis in the US began, in early 2008, NBRM initiated a project with the purpose to rigorously analyze movements in house prices in Macedonia. Since the data on house prices that were available from the State Statistical Office of the Republic of Macedonia (SSO) were not satisfactory (they were semi-annual, starting only from 2004, and were referring only to newly-built dwellings), it was decided to construct an own index. The index that was constructed was quarterly, starting from 2000, included both new and old dwellings, and was hedonic in nature, i.e. controlled for changes in quality. More details about the construction of the index can be found in Davidovska Stojanova et al. (2008). Below we provide only the most important aspects.

The index is constructed from advertised data, which are collected from newspapers. Usually, data are taken from the real estate agencies that advertise in the newspapers. The index refers only to the capital city, Skopje, as there are not many advertisements for other towns, and covers only apartments, not houses, for the same reason. Approximately 150-200 apartments are taken in each quarter. Usually, all zones/municipalities are covered. The index is published on the website of NBRM and is commented in the Quarterly Report (the flagship report of NBRM).

The index is calculated using the characteristics price method. More precisely, the following hedonic regression is estimated for every quarter:

$$\begin{aligned} price_i = & c + p_1 * area_i + p_2 * area_i^2 + p_3 * floor123_i + p_4 * floor89_i + p_5 * floor10plus_i \\ & + p_6 * central\_heating_i + p_7 * zone2 + p_8 * zone3 + p_9 * zone4 + p_{10} * zone5 \\ & + p_{11} * newly\_built_i + \epsilon_i \end{aligned} \quad (1)$$

where  $i$  indexes the apartments,  $\epsilon_i$  is the error term,  $p$ 's are the regression coefficients, and the names of the variables should be self-explanatory. The floor and zone variables are dummy variables ( $floor123$  being a dummy for apartments on the first, second or third floor,  $zone2$  a dummy for apartments from zone number 2, etc.), as well as the variables for central heating and newly built apartments.

The index is then calculated according to the following index formula:

$$index = \frac{\sum_{i=1}^n \sum_{j=1}^{11} p_{j,t+1} * q_{j,i}}{\sum_{i=1}^n \sum_{j=1}^{11} p_{j,t} * q_{j,i}} \quad (2)$$

where  $t$  indexes the time,  $i$  indexes the apartments, whose total number is  $n$ ,  $j$  indexes the characteristics of the apartments,  $q$  are the characteristics, shown in equation 1 (eleven in total), and  $p$  are the prices of the characteristics, i.e. the coefficients from hedonic regression.

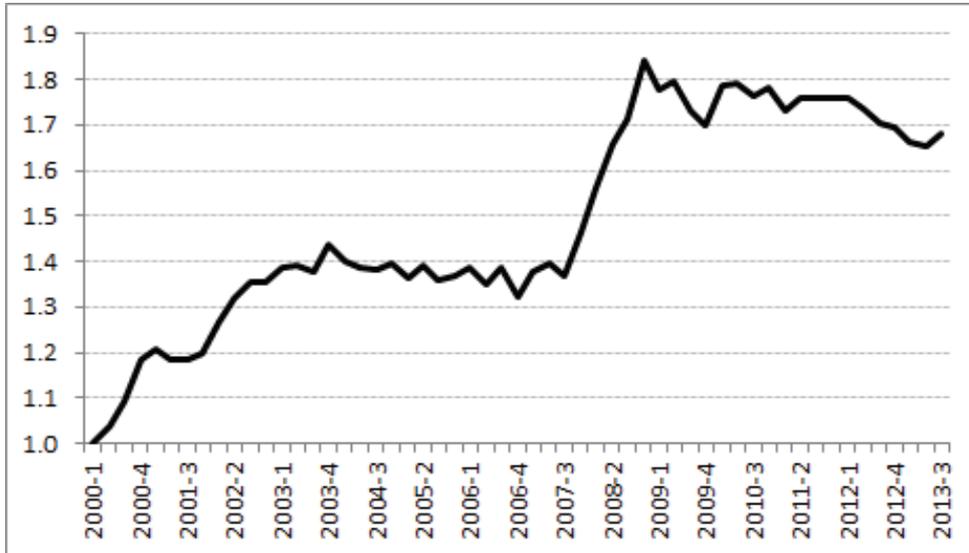
This is the standard index formula, with one small difference - that all the apartments available in the dataset are taken for calculating the price change between two periods, not just the apartments from the corresponding periods. This essentially implies that the apartments that are sold in each quarter are same with the total housing stock. While this is clearly not usual in the literature, we decided to do this in order to avoid biases emerging from

inappropriate selection of the base year.

Despite the relatively low number of explanatory variables in the hedonic regression, the regression fit is relatively good, with adjusted  $R^2$  above 90% in most of the periods.

The hedonic house price index is shown on figure 1.

FIGURE 1: THE HEDONIC HOUSE PRICE INDEX (2000Q1=1)



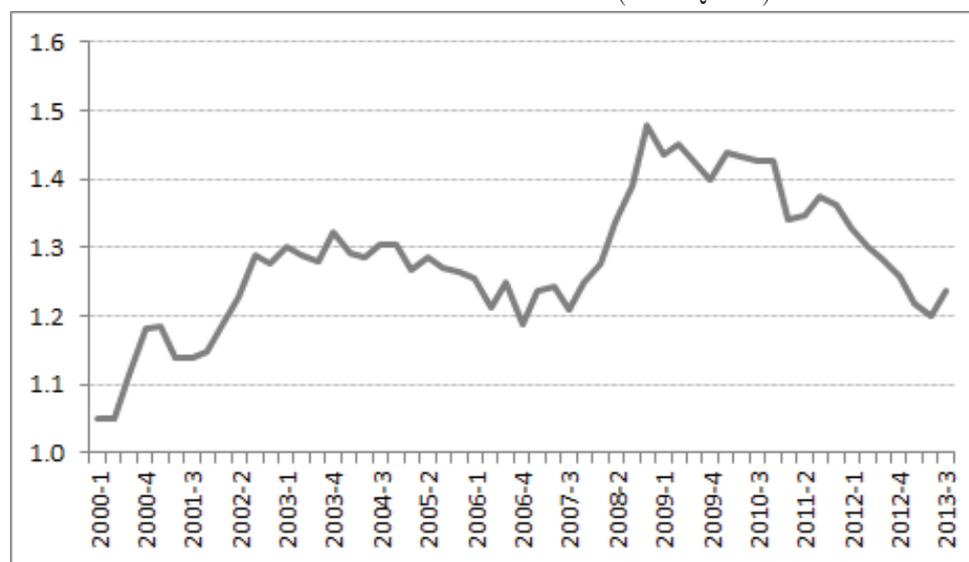
Source: NBRM

#### IV. GENERAL MOVEMENTS IN THE INDEX

Several things can be observed from Figure 1. First, the current value of the index is approximately 70 percent higher than the initial value, implying that the average annual growth rate of house prices in Macedonia during this 14-year period has been around 4 percent, which is not too much, especially bearing in mind that the average inflation in the corresponding period was approximately 2.5 percent. Second, there are clear cyclical movements in the index, consisting of periods of expansion and periods of stagnation/fall, each phase lasting around three years. The first expansion was from 2000 until end of 2003, when house prices rose by 44 percent in total, or approximately 10 percent per year. This expansion was followed by a general stagnation in the price for the next three-and-a-half years (until 2007Q3). The price started to rise again at the end of 2007 and increased by 35 percent until the end of 2008. Since then, house prices have been on a downward trend for five years, currently being 9 percent below the peak from 2008Q4. Third, it seems that house prices are rather rigid downwards - there was not an obvious fall in the price after the first expansion, while the decline after the second expansion has brought current prices at the level of 2008Q2 (i.e. only two quarters before the peak).

The last point does not hold when the real house price is observed (i.e. the house price index, divided by the consumer price index). As can be seen on Figure 2, real house prices have fallen rather substantially during the last five years - by 17 percent (approximately 4 percent per annum). We next try to explain what has caused this fall in the house prices.

FIGURE 2: REAL HOUSE PRICES (2000Q1=1)



Source: Author's calculations, using data from NBRM and SSO

## V. ECONOMETRIC ANALYSIS OF MACEDONIAN HOUSE PRICES

### V.A. *The model*

The model that we use to explain movements in (real) house prices is similar to that of Davidovska-Stojanova et al. (2008). It is essentially an operationalization of a standard small structural model of the housing market, where the equilibrium house price is derived from the housing demand and supply functions. More precisely, the demand is defined as:

$$D = f(\text{price}, \text{income}, \text{mortgages}) \quad (3)$$

The supply is defined as:

$$S = f(\text{price}, \text{newly\_built\_apartments}) \quad (4)$$

In equilibrium, demand equals supply, which implies that the equilibrium price is:

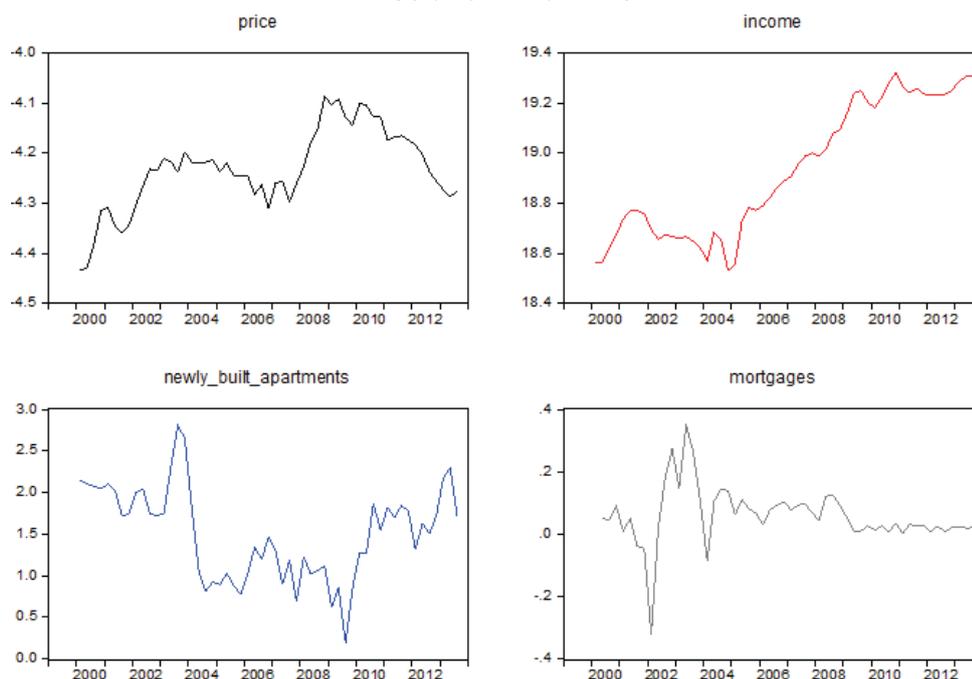
$$\text{price}_{\text{equilibrium}} = f(\text{income}, \text{mortgages}, \text{newly\_built\_apartments}) \quad (5)$$

Econometric techniques are used to estimate the above relationship (replacing the equilibrium price with the observed), and then the fitted values from the regression give the equilibrium price.

### V.B. *Variables and data*

All the variables from Eq. 5 are in real terms (i.e. divided by the consumer price index) and seasonally adjusted (using the Census X-12 method). Mortgages are in their dlog form (i.e. the quarter-on-quarter growth rate), since they are a stock variable. All the other variables are in levels (logs of the levels). The income is approximated as the average nominal wage, multiplied with the number of employed persons. The source of all data is SSO, except the house price index, which is from NBRM. The variables that enter the model are presented on Figure 3.

FIGURE 3: VARIABLES



Source: Author's calculations, using data from NBRM and SSO.

The variables seem non-stationary. To formally test for this, we apply the test proposed by Ng and Perron (2001). This test has better size and power than the early unit root tests, such as Augmented Dickey-Fuller, or the Phillips-Perron test. The null hypothesis of the test is that the series has a unit root, and test values below the critical values indicate that the null can be rejected (i.e. the series can be treated as stationary). In all the tests, we allow for both trend and intercept. The transformation of the variables is done using the Autoregressive Generalized Least Squares detrending, and the number of lags in the regression is determined on the grounds of the Modified Akaike criterion (with maximum number of lags equal to 10). We present the four statistics proposed by Ng and Perron (2001),  $MZ_{\alpha}^d$ ,  $MZ_t^d$ ,  $MSB^d$ ,  $MP_T^d$ , in Table 1.

TABLE 1: UNIT ROOT TESTS

Variable	$MZ_{\alpha}^d$		$MZ_t^d$		$MSB^d$		$MP_T^d$	
	Test stat.	5% crit. val.	Test stat.	5% crit. val.	Test stat.	5% crit. val.	Test stat.	5% crit. val.
price	-2.9	-17.3	-0.96	-2.9	0.33	0.17	24.8	5.5
income	-4	-17.3	-1.4	-2.9	0.35	0.17	22.7	5.5
mortgages	-25.1	-17.3	-3.5	-2.9	0.14	0.17	3.6	5.5
newly_built_apartments	-3.6	-17.3	-1.3	-2.9	0.35	0.17	24	5.5

Price, income and newly built apartments seem non-stationary, since the test statistics is bigger than the critical value. Their first differences seem stationary (results available upon request), hence these series can be treated as integrated of order 1. On the other hand, mortgages seem stationary.

### *V.C. Econometric analysis*

Most of the variables appear to be non-stationary, so the econometric analysis will be done using cointegration techniques. We apply the Vector Error Correction method (VECM), developed by Johansen (1988, 1991). Although mortgages appear to be stationary, we still decide to include them in the cointegrating vector, as we believe they are very important for the equilibrium price.

The Johansen techniques starts with a Vector Autoregression (VAR), which is later transformed in a VECM. The first step of the analysis is then to decide on the number of lags in the VAR. We decide this on the grounds of five information criteria. Table 2 presents the number of lags suggested by the five criteria, with \* indicating the number of lags suggested by the corresponding criterion.

TABLE 2: VAR LAG ORDER SELECTION CRITERIA

Lag	LR	FPE	AIC	SC	HQ
1	NA	1.85E-09	-8.75	-8.15*	-8.53*
2	10.6	2.75E-09	-8.37	-7.15	-7.91
3	45.4*	1.62e-09*	-8.93*	-7.09	-8.23
4	20.2	1.79E-09	-8.88	-6.43	-7.95

\* indicates lag order selected by the criterion

LR: sequential modified LR test statistic (each test at 5% level)

FPE: Final prediction error

AIC: Akaike information criterion

SC: Schwarz information criterion

HQ: Hannan-Quinn information criterion

Three of the five criteria (LR, FPE and AIC) suggest including 3 lags in the VAR, while SC and HQ suggest including only 1 lag. We decide to proceed with 3 lags, also because 1 lag seems to be insufficient, a priori, for capturing the dynamics between the variables.

We next test the residuals from the 3-lag VAR for serial correlation, heteroskedasticity and non-normality. Table 3 presents these results. The null hypotheses of no serial correlation and no heteroskedasticity in the residuals cannot be rejected at the 1% level of significance, while the null hypothesis of normality can be rejected. Despite the likely non-normality, we proceed with this VAR.

TABLE 3: VAR RESIDUALS DIAGNOSTICS TESTS

<i>Serial correlation LM test</i>	
<b>Null hypothesis</b>	<b>P-value</b>
No serial correlation of order 1	0.04
No serial correlation of order 2	0.07
No serial correlation of order 3	0.54
No serial correlation of order 4	0.01

<i>White heteroskedasticity test</i>	
<b>Null hypothesis</b>	<b>P-value</b>
Homoskedasticity in the residuals	0.08

<i>Jarque-Bera normality test</i>	
<b>Null hypothesis</b>	<b>P-value</b>
Residuals are normally distributed	0.00

We next test for cointegration, using the trace and maximum eigenvalue statistics. The results of the test are presented in Table 4.

TABLE 4: COINTEGRATION TESTS

	<b>Option 1</b>	<b>Option 2</b>	<b>Option 3</b>	<b>Option 4</b>	<b>Option 5</b>
	<i>No trend, no intercept</i>	<i>No trend, intercept in CV</i>	<i>No trend, intercept in VAR</i>	<i>Trend in CV, intercept in VAR</i>	<i>Trend and intercept in VAR</i>
Trace	1	1	1	1	1
Max-Eig	1	1	1	1	1

The table shows the number of cointegrating vectors suggested by the two tests, trace and maximum eigenvalue, on the 5% level of significance, for the 5 options.

Critical values based on MacKinnon, Haug and Michelis (1999).

CV stands for 'cointegrating vector'.

The results suggest that there is likely one cointegrating vector according to the both tests, in all the five combinations of the deterministic elements (intercept and trend). We decide to work with option 3, the most reasonable option. We discard option 4 because we do not believe there should be a deterministic trend in the house prices.

We normalize the cointegrating vector on the house prices (i.e. we set house prices as the dependent variable in the cointegrating vector). The cointegrating vector, transformed as an equation, is presented below (Eq. 6). We do not report the short-run equations, to conserve space, but they are available upon request.

$$price = 0.17*income + 0.57*mortgages - 0.18*newly\_built\_apartments$$

(2.8)                      (6.6)                      (2.9)                      (6)

Absolute t values in parentheses. Values above 2 can be loosely treated as significant at 5%.

The interpretation of the vector is as follows. If real income increases by 1%, real house price will increase by 0.17%, in the long run, on average, *ceteris paribus*. If the quarter-on-quarter rate of growth of real mortgages is higher by 1 percentage point, real house price will be higher by 0.57 percent, in the long run, on average, *ceteris paribus*. If the real value of the newly built apartments increases by 1 percent, real house will fall by 0.18 percent, in the long run, on average, *ceteris paribus*.

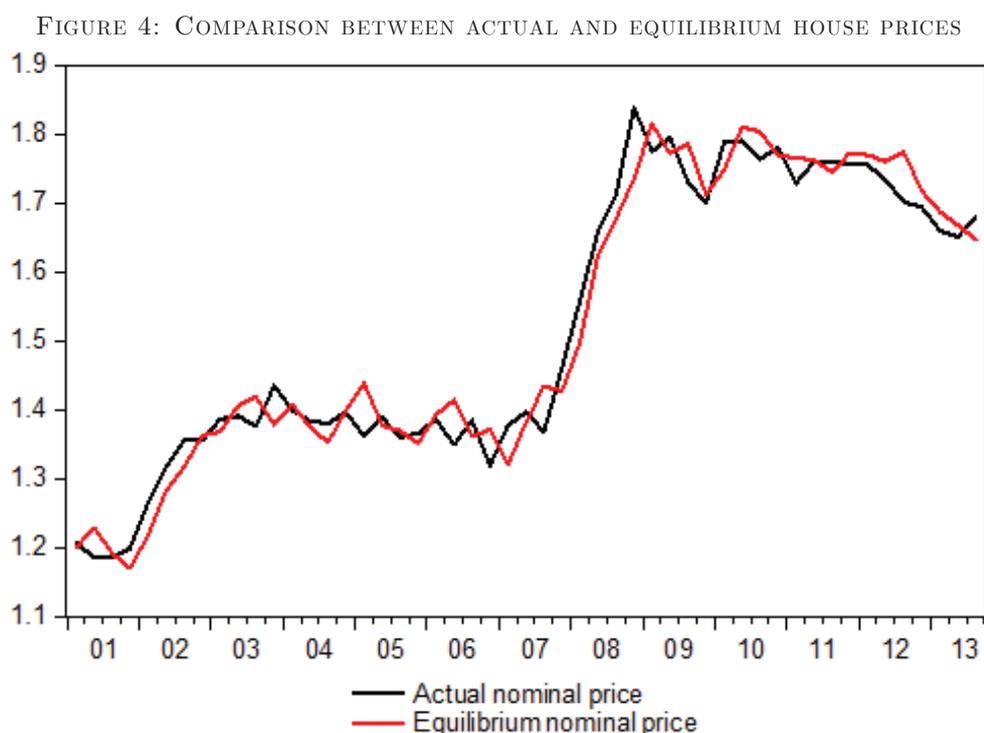
#### ***V.D. Explaining recent movements in house prices***

The above results can be used to explain the recent movements in the house prices. The real house price fell substantially during the crisis - the index in 2013Q3 is approximately 17 percent lower than the peak reached in 2008Q4. These movements are entirely in accordance with the movements in the underlying fundamentals. Real income has been stagnating since 2009 - the average annualized quarterly growth rate during this period has been around 1.5 percent, pushing the price up by only 0.3 percent per year. Mortgage growth has slowed down, from approximately 8.5 percent per quarter in 2006-2008, to approximately 2 percent per quarter in 2009-2013, pushing real house price down by around 4 percent. Finally, the supply of new apartments in 2010-2013 has been higher by approximately 60 percent than in 2007-2008, which has pushed house prices down by approximately 11 percent. All in all, the decline in the real house prices that has been going on for the last five years is a consequence of the lower demand for apartments, due to the stagnation in real income and the slow down in the credit activity, and the higher supply of new apartments.

How can one explain the surge in the supply of new apartments in Macedonia during the Great Recession? This may be a topic for a research on its own, but we still provide some insights. One potential explanation can be through the spill-overs from the higher government investment observed during this period. Namely, since 2008, there has been a surge in public investment in Macedonia - the average value of real general government capital expenditure in 2008-2013 has been nearly 50 percent higher than in the period 2005-2007. Big part of this investment is in public sector buildings located in the capital city (the so-called 'Skopje 2014' project). Hence, the surge in the newly-built apartments can be a consequence of the economies of scale and the lower fixed costs of construction of companies involved in the public investment projects. Some support for this explanation can be found in the notion that the supply of new apartments and the government capital expenditures seem to be cointegrated, with a coefficient of 0.9 (these results are not presented, but are available upon request).

### *V.E. Comparison between actual and equilibrium price*

The most important question for policy-makers is whether the actual house price has been aligned with the fundamental (or equilibrium) price. The very existence of cointegration between the price and the fundamentals implies a positive answer to the above question, but we still present a comparison between the actual and the equilibrium price. Here we present the *nominal* prices, on Figure 4.



*Source: Author's calculations, using data from NBRM.*

In general, the actual price has been fluctuating around the equilibrium price in the period 2000-2013, without pronounced and persistent misalignments. The highest misalignment has been observed in 2008Q4 (the period when the price reached its peak), when the actual price was above the equilibrium one by approximately 6%. Three periods of somewhat persistent, though small, misalignments, can be identified - 2001Q4-2002Q3, 2007Q4-2008Q4 and 2011Q4-2013Q2. During the first two periods, the actual price was above the equilibrium one, and was rising. During the third period, the actual price was below the equilibrium one, and was declining. During the most recent eight quarters (2011Q4-2013Q3), the actual price has been below the equilibrium price by 1 percent, on average, which is modest, by any standards.

---

## VI. CONCLUSION

This paper presents an overview of developments in the house prices in Macedonia in 2000-2013. During this period, house prices in Macedonia have exhibited a typical cyclical behaviour, with periods of growth followed by periods of stagnation/fall. These movements have been in accordance with the fundamentals, like income, mortgages and supply of new housing, and during this period apartments have not shown pronounced and persistent misalignments from the fundamental (equilibrium) price. Since the onset of the Great Recession, real house prices in Macedonia have declined by 17 percent, but this decline is again in accordance with the fundamentals, i.e. with lower demand during this period, due to the stagnation in real income and the slowdown in the credit growth, and the higher supply of new housing.

As a final remark, the methodology presented in this paper is used more or less regularly at the NBRM for assessing whether house prices are overvalued, and has shown satisfactory real-time performance. Similar analysis to this one has been published in the April 2009 Quarterly Report, which was written when house prices were at their peak. It was reported that "*[t]he overpricing in the fourth quarter of 2008 equals 6%*" (NBRM, 2009, p.29), and that "*regarding the probable future movement of the price of the apartments, it can be stated that the fact that in the last three quarters the apartments show signs of overpricing indicates possible downward correction in the price in near future*" (NBRM, 2009, p.30). The overvaluation from 2008Q4 is still assessed at 6 percent, and house prices indeed started declining in 2009.

## VII. REFERENCES

DAVIDOVSKA-STOJANOVA, BILJANA, BRANIMIR JOVANOVIĆ, MAJA KADIEVSKA-VOJNOVIĆ, MAGDALENA PETROVSKA AND GANI RAMADANI (2008). "Real Estate Prices in the Republic of Macedonia," National Bank of the Republic of Macedonia Working Paper, August 2008.

JOHANSEN, SOREN (1988). "Statistical analysis of cointegration vectors," *Journal of Economic Dynamics and Control*, Elsevier, Vol. 12, No. 2-3, pp. 231-254.

JOHANSEN, SOREN (1991). "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, Vol. 59, No. 6, pp. 1551-1580.

MACKINNON, JAMES G., ALFRED HAUG, AND LEO MICHELIS (1999). "Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration", *Journal of Applied Econometrics*, Vol. 14, No. 5, pp. 563-577.

NBRM, NATIONAL BANK OF THE REPUBLIC OF MACEDONIA (2009). *Quarterly Report, April, 2009*. Skopje, Macedonia: National Bank of the Republic of Macedonia.

NG, SERENA AND PIERRE PERRON (2001). "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica*, Vol. 69, No. 6, pp. 1519-1554.

STIGLITZ, JOSEPH E. (1990). "Symposium on Bubbles," *Journal of Economic Perspectives*, Vol. 4, No. 2, pp. 13-18.

---

## **17. Impact of the economic crisis on house prices in the Czech Republic measured on hedonic price index on bank data**

**Martin Lux, Petr Sunega**



Martin Lux, Petr Sunega – Department of Socioeconomics of Housing, Institute of Sociology,  
Academy of Sciences of the Czech Republic, Jilská 1, 110 00 Prague 1, martin.lux@soc.cas.cz,  
petr.sunega@soc.cas.cz

This paper was prepared under the Research Project sponsored by the Czech Science  
Foundation with grant number P404/12/1446.

---

## Article 17 Contents

Abstract .....	176
Introduction .....	177
The hedonic house price index in the Czech Republic .....	188
Analysis of the impact of the economic crisis on housing and mortgage markets .....	190
Conclusions .....	198
References .....	201

## **Abstract**

In the first section, this paper describes the methods employed to calculate the first ever hedonic price index created in the Czech Republic using data from the country's biggest mortgage lenders. This section also presents the wider context of the creation of this index and its practical application. In the second section, hedonic price index data are used to demonstrate the impact of the global economic crisis on housing prices in the Czech Republic. The impact of the economic crisis on housing and mortgage markets is then analysed using methods from both neoclassic and institutional economics. The findings show that institutional factors relating to the housing system may have an important influence on the scope and character of the impact of the economic crisis on housing and mortgage markets in the Czech Republic.

**Keywords:** hedonic price index – Czech Republic – house prices – global financial crisis

**JEL codes:** G12, G21, E32

## Introduction

This paper briefly describes, in the first section, the background and the methodology for calculating the hedonic house price index based on data from Czech mortgage lenders. This price index can be used, among other things, to observe the current trends in house prices. Consequently, unlike other price indexes available in the Czech Republic, it is also suitable for measuring the current impact of the economic crisis on house prices. In the second section, the paper focuses on the actual effects of the global economic crisis on the housing and mortgage markets in the Czech Republic. The impact of the economic crisis on housing and mortgage markets is then analysed using methods from both neoclassic and institutional economics. The findings show that institutional factors relating to the housing system may have an important influence on the scope and character of the impact of the economic crisis on housing and mortgage markets in the Czech Republic.

## The hedonic house price index in the Czech Republic

Several price indices on residential real estate have emerged in the Czech Republic since the late 1990s. Leaving aside the amateur attempts of some real estate agencies, which use simple averages of offer prices, there are three indexes that from a methodological perspective can be regarded as relevant. The first is the index published by the Czech Statistical Office (CSO) each year in the publication *The Prices of Monitored Types of Real Estate*, which is based on transaction prices ascertained by financial offices and recorded by the Ministry of Finance for the purpose of stamp duty collection. The second is the index created by the Institute of Regional Information (IRI), which is based on offer (advertised) prices. The third index – the price index of the Czech Technical University (CTU) – is saturated using a similar method (taking offer prices from adverts).

All of the above-cited indexes grapple with methodological problems, relating either to the indexing method and sample selection (IRI, CTU) or to the scope and quality of the data (CSO). In the data file for the CSO index the sample of executed transactions is incomplete, only a small number of housing attributes are collected, and above all the index itself is released with a substantial delay (by as much as two years). The approaches used by the IRI and the CTU are based on advertised (offer) prices; data on real estate attributes depends on how truthful and exhaustive the information in the adverts is. Moreover, for the purpose of the index only ‘standard’ flats are used; the hedonic approach is not applied.

The hedonic approach to measuring changes in house prices views the price of a real estate unit as the sum of the implicit (‘concealed’) prices of that real estate’s attributes (for example, the size of the flat, its age, the attractiveness of its location, and other attributes) and in this way it offers the best examination of residential real estate as a complex economic good. The following equation articulates the hedonic price model in its most general form:

$Y = B_i S + B_j L + B_k R + \varepsilon$ , where:

- $Y$  - price of the real estate unit;
- $S$  - attributes of the given real estate unit (flat size, number of rooms, etc.);
- $L$  - location-specific attributes (the socioeconomic features of the location, accessibility to the workplace, etc.);
- $R$  - region-specific attributes (the level of the region's economic development, employment rate, etc.);
- $\varepsilon$  - model error component (residuals).

The hedonic approach involves identifying the relationship between price-determining factors (the real estate's attributes) and price. This relationship is first tested for each component separately, that is, the relationship between the dependent variable and each explanatory variable, and only then are the multi-dimensional models capturing the interactive effect of all the explanatory variables on the dependent variable tested. If the relationships between the dependent and the explanatory variables are non-linear, a mathematical transformation is employed. Many explanatory variables enter the model, for instance, in quadratic, cubic, or logarithmic form or after a so-called gamma transformation, and even the functional form of the entire model is changed, usually by means of a semi-logarithmic transformation, where the dependent variable of 'house price' is replaced with the 'natural log of the housing price'.

The advantage of using hedonic models to construct a house price index is that they are relatively best able to identify the net price change and control for the changes in the quality and structure of transactions occurring during the period to which the index pertains. Almost all the data on transactions can be used; this can be contrasted with the repeat sales index, which uses information on around just 20% of transactions on houses repeatedly sold. The main disadvantage of the hedonic index is the volume of observed data on individual real estate units that has to be

monitored. The more thoroughly price changes are to be cleaned of changes in the quality and structure of units sold, the more detailed the information on the attributes of real estate units that has to be monitored.

A good way of obtaining reliable and comprehensive data records on real estate units and their close surrounding is to employ independent real estate appraisers, who make assessment of real estate for mortgage applications. Real estate appraisers are in a relatively independent position; their remuneration is not tied to the client's interests. Unlike the information on the attributes of a real estate unit provided by real estate agencies, which have a stake in the transaction, the information recorded by an appraiser can be regarded as more reliable; often this information is verified by an internal auditing system of mortgage lender. The standardised, electronic form of records they use ensures that the records are complete - all the important fields in the appraisal form must be filled in, otherwise the appraiser cannot electronically submit the appraisal for processing. In real estate adverts it is common for there to be missing information on the various attributes of the real estate unit.

For this reason, the source of data selected to construct the first hedonic price index in the Czech Republic was the database of residential real estate pledged to mortgage lenders operating in the Czech Republic. The objective of the project to construct this index was to create a dataset shared and saturated by all or a large portion of mortgage lenders in the Czech Republic. Bank data-sharing results in statistical models of much greater reliability, leads to more detailed monitoring of price trends in different market segments, eliminates the deviation caused by differences in the transaction portfolios of different banks, and generates lower unit costs.

However, to achieve this objective, it was first necessary to establish a uniform format for electronic data collection that will be applied by all mortgage lenders, and this proved to be a very time-demanding process. In cooperation with Česká spořitelna, a.s., the second-largest mortgage lender in the Czech Republic (part of

ERSTE financial group), in 2005 a new format for data collection was devised, and, at no small expense to the bank, which had to change its internal IT system, it began to apply this data collection form in 2007. After testing the hedonic models on data collected in 2007 and 2008, the number of relevant attributes in the data collection form was reduced to the 78 most significant attributes.

In the spring of 2008 this form was also adopted by representatives of Hypoteční banka, a.s., the largest mortgage lender in the Czech Republic (part of KBC financial group). These two banks together hold approximately 50% of the total outstanding mortgage balance in the Czech Republic. In 2010 the price index was calculated for the first time using the shared data of these two biggest mortgage lenders in the Czech Republic. In 2013 two additional mortgage lenders entered the project: UniCredit Bank and Wüstenrot. The coverage of market by data used for index on the shared data increased to approximately 60% of outstanding mortgage balance.

Other mortgage lenders are being invited into (or being informed about) the project on an ongoing basis through a working group set up within the Czech Banking Association. Despite our repeated requests the project received no financial support from the Czech National Bank, even though both it and the Ministry of Finance voiced their support for the project.<sup>1</sup>

Given that real estate appraisers do not have at their disposal all the information on a given location or its accessibility, to calculate the hedonic models the data provided by appraisers and collected by mortgage lenders are supplemented with additional and regularly updated data on location and region specific attributes. The following data were supplemented to construct the hedonic models: selected data from the Population and Housing Census for each cadastre district based on sources from the

---

<sup>1</sup> The costs incurred in connection with the establishing a uniform data collection format are mainly motivated by the need to revalue pledged real estate at regular intervals – and better and more reliable revaluations can be attained by sharing data. The resulting dataset is thus not used just for the price index, but also, and by the mortgage lenders especially, for automatic revaluations of pledged real estate.

Czech Statistical Office (percentage of the population with different levels of education; age structure of the population; the share of the housing stock built using panel technology); data on the average unemployment rate for each municipality (based on sources from the Ministry of Labour and Social Affairs); data on the average monthly number of recipients of means-tested social benefits (specifically recipients of the housing allowance and the child allowance) for each municipality (based on sources from the Ministry of Labour and Social Affairs); data on the population size for each municipality (based on sources of the Czech Statistical Office); data on the services and facilities available in municipalities based on the database of Statistics on Towns and Municipalities maintained by the Czech Statistical Office for each municipality (net migration, utilities the municipality is connected to, the range of cultural, educational, health and sporting facilities, the scope of housing development); and data on the number of self-employed for each municipality (based on sources from the Czech Statistical Office's Register of Economic Entities).

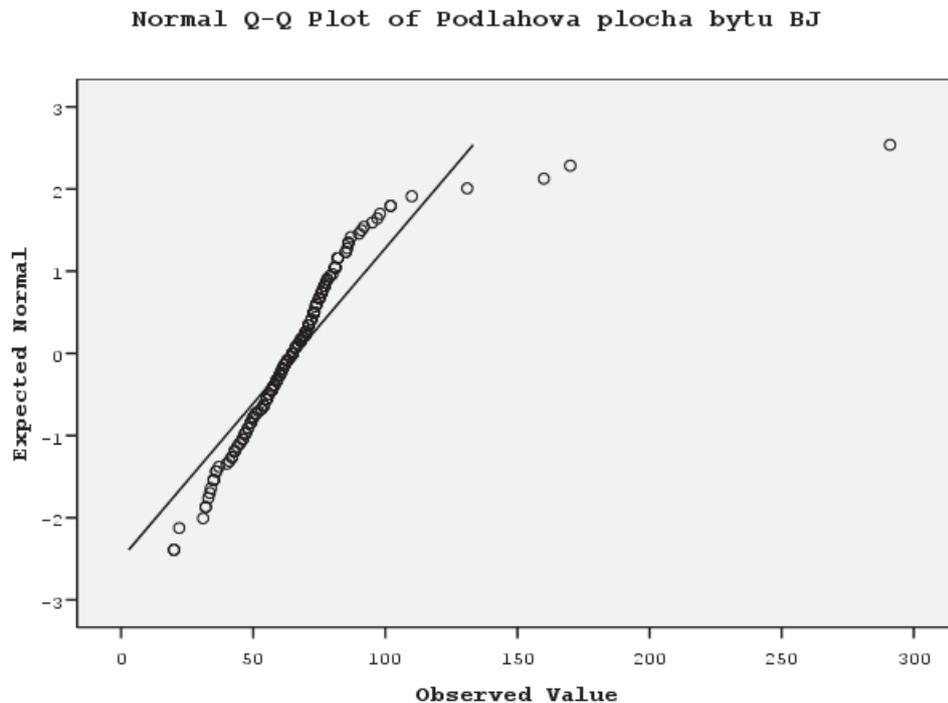
Using an ID number (the code of the cadastre or the code of the municipality) these external data are 'automatically' assorted into a data file on the prices and attributes of real estate units. Also included in the data file are the commuting distances from the municipality in which the real estate unit is located to regional capitals and to 'centres of employment' – the fastest commuting route (in kilometres) and the fastest commuting time (in minutes) to the closest 'centre of employment' and the closest regional capital. 'Centres of employment' are defined as municipalities with more than 15,000 inhabitants. Technically the above-mentioned commuting distances were obtained from publicly accessible map websites using a script generating departure and destination points.

In order to prevent the systematic deviation of the results of hedonic regression models due to the inclusion of cases with extremely high or extremely low values (so-called 'outliers'), a group of key variables is identified for each of the individual segments of residential real estate (flats, family homes, land) and also for the

individual regions of the Czech Republic, and using these variables real estate units with extreme values are mechanically excluded from dataset. For flats the key variables are usually the size of the flat (floor area in  $\text{m}^2$ ), unit price (price per  $\text{m}^2$  of floor area), and total flat price.

Figure 1 provides a clear illustration of the approach used to establish the limits for excluding extreme values and at the same time ensuring that one of the basic assumptions of regression modelling is met: that exogenous variables have normal distributions. The observed values of the relevant indicator (in this case the floor area of the flat) are plotted in relation to the axis of the ideal normal distribution. Figure 1 shows that flats with a floor area greater than approximately  $140 \text{ m}^2$  are much further from the size of floor area of the other flats in this segment, so in this case a floor area of  $140 \text{ m}^2$  was selected as the maximum floor area for flats included in the regression model.

**Figure 1: Mechanical filtering – the Central Bohemia Region, flats, municipalities with a population of up to 4,999 inhabitants**



After excluding real estate units with extreme or irregular values, supplementary use is made of a statistical filtering method. Statistical filtering is based on an analysis of residuals, i.e. the differences between the observed values of the dependent variable (estimated prices or their logarithm) and values predicted by the model. The regression model is calculated with the aim of identifying outlying observations, and these observations are not entered into the calculation of the regression model in its final form in the second step. Whether the observations are outliers is determined on the basis of the size of the deviation of the real price of the flat and the price predicted by the model. This is a standardised residual (*Studentized residual*), which is determined as the absolute value of the difference between the estimated price and the price predicted by the model divided by the estimated standard deviation of the residual for the given flat. In order to identify extreme and outlying observations, a maximum 'cut-off' level for the absolute value of the residual has to be established. No general recommendations exist on how to establish this cut-off level, so it is determined empirically in reference to the number of excluded real estate units (the cut-off rate must not be too high) and in reference to the results of the model (the explanatory and predictive power must be satisfactory).

The hedonic price index is constructed with data that have been cleaned from outliers and supplemented as described above. The hedonic price model is computed for the Czech Republic as a whole (the regions are represented just in the form of input dummy variables) and it is maintained in terms of the structure throughout all time periods; only parameters are regularly recalculated. Depending on the length of the period under observation, a price index can be monthly, quarterly, semi-annually, and so on; for the time being, a quarterly index was applied on mortgage lenders data in the Czech Republic.

There are basically three ways in which the factor of time can be incorporated into hedonic price models and the price index can be built. The easiest way, but also the one most criticised, is to use the so-called pooled time-dummy method. The

principle of this method involves calculating a single regression equation from the shared data file for multiple observed periods. Here, alongside the individual attributes of real estate units and attributes of location, time is included in the equation in the form of dummy variable. This approach, however, implicitly assumes that the change in prices over time is manifested in equal proportions across all the real estate attributes in the equation and thus that there is no change in the implicit prices of individual attributes over time.

The second approach is the so-called adjacent-period time-dummy method. The principle of this method involves calculating the regression equation from the shared data file for just two successive periods. The implicit prices of individual real estate attributes would be different for each of the equations, i.e. there would be some relaxation of the rigid assumption that the implicit prices of individual attributes experience no change. That rigidity, however, is only relaxed to some degree. The third and by far most flexible approach is the so-called hedonic imputations method. Using this method, as many regression models as there are periods for which a price change is to be determined are calculated. In this case there is no dummy variable representing time in the regression equations and the restriction on changes to the implicit prices of the individual real estate attributes is entirely eliminated. The equation coefficients, that is to say, the implicit prices of individual attributes, can differ over time.

Generally it could be said that the advantage of the first two methods using a time dummy variable derives mainly from the fact that the regression models are calculated from shared data files, i.e. from the largest sample of real estate units, and this helps make the parameters of the estimated models more robust. The disadvantages have already been discussed: the main disadvantage remains the rigidity attached to determining the implicit prices of individual attributes, and the lack of adaptability to structural changes in supply and demand in the residential real estate market that could occur in the future.

For the reasons cited above, the hedonic imputations method was used to create a hedonic price index based on the data from mortgage lenders in the Czech Republic. However, in such a case it is necessary to ensure that potential changes in the structure of observed real estate units between periods are adequately controlled for and this is done with the help of weighting (mix adjustment). In other words, before finally comparing average real estate prices, it is first necessary to weight house prices for each combination of real estate attributes. Usually these weights are calculated from the moving averages of the total number of transactions in different segments over the past several years. These can be derived either from the data of the mortgage lenders or from the data of the cadaster offices. If no such aggregated data exist – which was the case in the Czech Republic – it is possible to use census data. The disadvantage of using the census is generally that it is impossible to adjust the weights continuously to the changing structure of the housing stock over time. However, because no other option was available, data from the 2001 census were firstly used to create the weights. In 2013, i.e. after three years of operating the index on the shared data, transaction weights were computed on sample of all transactions monitored between 2010 and 2012 and the transaction weights are going to substitute census-based weights in the last quarter of 2013. Transaction weights will be recomputed each year in the form of a three years moving average.

Compared to the Czech Statistical Office price index, the hedonic price index based on data from mortgage lenders offers:

- *Up-to-date information* – the index is calculated quarterly and the results of the calculation are usually released within three weeks of the end of the given quarter. The CSO index can never attain this standard. The CSO transaction price index has more than a one-year delay, and even after that it is substantially adjusted owing to later additions to the dataset;
- *A vast dataset with which to create the index* – the index is created on detailed data on the individual attributes of real estate that are obtained directly from independent appraisers based on their viewing of the real

estate. The dataset itself contains large number of real estate attributes, which makes it possible to reliably apply the hedonic approach. The CSO index contains very few attributes and the application of the hedonic approach is limited;

- *A high quality of data, and an index based on appraised (real) and not offer prices* – the dataset relates to real transactions and is created by independent appraisers; it does not relate to offer prices or information from developers or real estate agencies;
- *Sustainability of the index methodology* – the data collection is done using methods that should basically remain unchanged in the future (the principles of mortgage lending, real estate appraisals). On the other hand, the stamp duty could be eliminated in the future for political reasons – then the dataset required to create the CSO index would no longer exist.

Figure 2 shows the price index on the shared data (data of all mortgage lenders participating in the project) for different residential market segments: apartments (flats), detached houses and building plots. The figure clearly shows the price drop for apartments, price stagnation for detached houses and price increase for building plots during preceding three years. However, given that the second mortgage lender (Hypoteční banka) did not join the project until 2010, the results of the hedonic price index for apartments based only on the data of Česká spořitelna are also presented (Figure 3). With this time series it is possible to analyse the impact of the global economic crisis on prices of apartments in the Czech Republic because the co-operation with Česká spořitelna, the second largest mortgage lender, has started already in 2007.

Table 1 presents the fit of the hedonic regression price models used for computation of Česká spořitelna index: the adjusted R-squared and average deviation between real and estimated price for all the hedonic models since the start of price observations (Q4 2007). The index results show the decrease in prices of flats between Q3 2008 (the index value of 110) and Q3 2013 (the index value of 90).

Compared to the situation in many other countries the decrease in apartment prices is not particularly dramatic. Furthermore, the prices of detached houses did not change and prices of building plots actually increased during this period.

**Figure 2: Hedonic price index on the shared data (Q1 2010 - Q3 2013)**

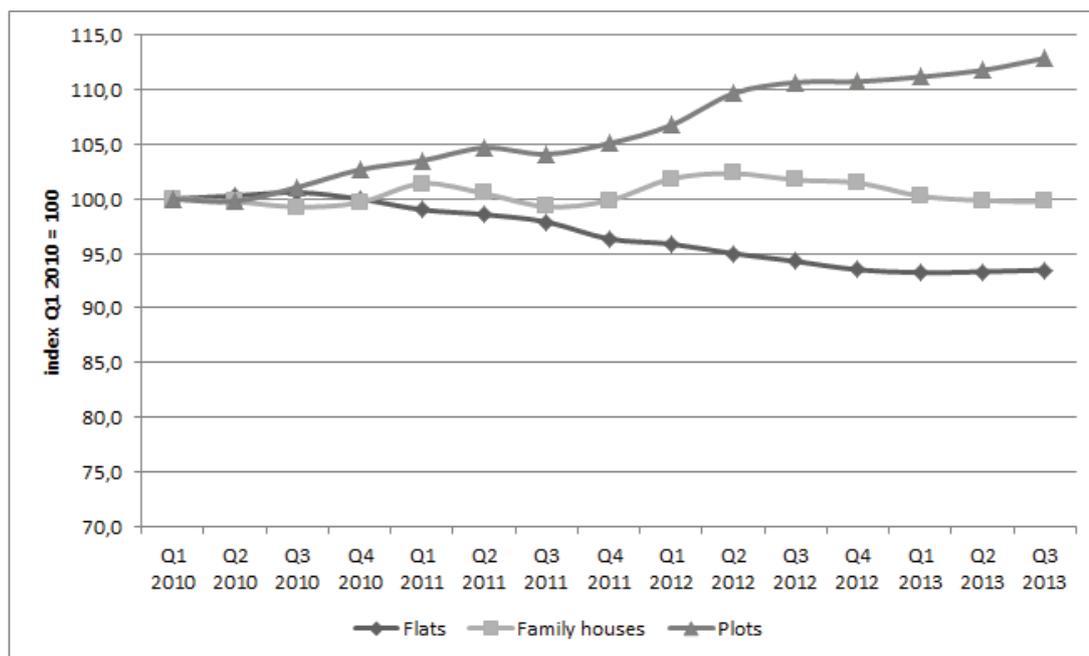


Figure 3: Hedonic price index for apartments, Česká spořitelna index (Q4 2007 – Q3 2013)

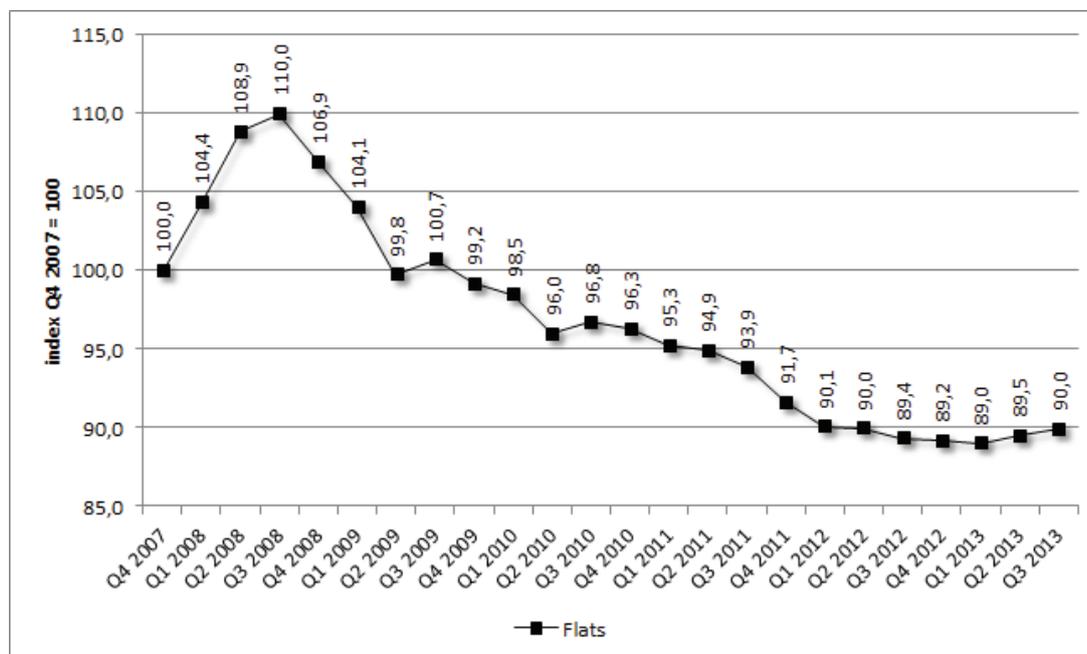


Table 1: The fit of the hedonic price models, apartments, Česká spořitelna index (Q4 2007 – Q3 2013)

Quarter	Adj. R2	Deviation (%)
Q4 2007	88.8	12.5
Q1 2008	90.1	11.8
Q2 2008	89.5	11.7
Q3 2008	89.9	12.0
Q4 2008	89.3	12.2
Q1 2009	87.4	12.1
Q2 2009	90.0	12.0
Q3 2009	89.5	12.1
Q4 2009	90.7	11.9
Q1 2010	89.7	11.2
Q2 2010	90.1	12.3
Q3 2010	90.6	11.6
Q4 2010	90.2	11.4
Q1 2011	88.3	12.7
Q2 2011	88.0	12.1
Q3 2011	89.9	12.1
Q4 2011	89.5	12.4
Q1 2012	90.4	12.5
Q2 2012	90.4	12.4
Q3 2012	89.4	12.4
Q4 2012	88.7	12.7
Q1 2013	88.5	14.2
Q2 2013	89.0	14.5
Q3 2013	90.4	13.8

## **Analysis of the impact of the economic crisis on housing and mortgage markets**

The second goal of this paper is to answer the central question: what factors explain the impact of the global economic crisis on mortgage and housing markets in the Czech Republic?

An analysis of markets and their reactions to external economic shocks may be framed in the terms of both neoclassical and institutional economic theory. Neoclassical economics applies the concepts of market efficiency and market equilibrium. Ideally, markets provide an optimal allocation of resources on the basis of consumer sovereignty. Differences in the performance of housing markets can be explained by differences in the size of the markets, the macroeconomic environment (such as inflation, the interest rate, unemployment), the existence of specific constraints (such as land or credit availability), market inefficiencies (such as monopolistic competition), and market regulations (such as the given legal framework). Consequently, the impact of the global economic crisis on mortgage and housing markets should be examined through an analysis of these ‘universal’ aspects of the economy and markets.

In contrast, institutional economists emphasise the importance of a specific institutional context, that is to say, the behaviour of organizational and individual actors embedded in specific cultural and social norms. Such norms are seen to be the product of decades of social interaction oriented towards solving the social, economic and political problems of a particular society (North 1991, Hodgson 2006). One of the key concepts of this theoretical approach is institutional equilibrium, which occurs when there is of no advantage to any market participants to take on the costs connected with altering existing contracts. Institutional equilibrium does not represent only ‘universal’ free market equilibrium. This is because in different historical and cultural contexts diverse institutional equilibriums will emerge. This is also the reason why the same global (exogenous) change will affect any pair of countries in different ways. In other words, the economic life of

---

markets is territorially embedded in social and cultural relations. Consequently, the observed impact of the recent global economic crisis on housing and mortgage markets should be explored through an analysis of the wider institutional and cultural context of Czech society.

Both neoclassical and institutional economics can enhance our understanding of market reactions and performance. Consequently, we can formulate two hypotheses that reflect the main causal mechanisms postulated by each of these economic theories, i.e. neoclassical and institutional economics, respectively.

H1 The impact of the global economic crisis on the mortgage and housing markets in the Czech Republic is explained by the macroeconomic environment, that is, for instance, by inflation, and the features of market regulations, such as the existence of regulations that expand or constrain the supply of mortgage credit.

H2 The impact of the global economic crisis on the mortgage and housing markets in the Czech Republic is explained by the wider institutional factors that influence the behaviour of market actors. Here the focus is on the institutional development of a housing system and housing tenure choice.

The impact of the economic crisis on the mortgage and housing markets itself will be measured by examining changes in house prices and the default rate among mortgage borrowers that emerged during the crisis period, i.e. between 2008 and 2012.

The Czech Republic's macroeconomic performance was generally very good before the economic crisis (2008). This is demonstrated by the fact that the Czech Republic maintained a high GDP growth with a relatively low public budget deficit and public debt. The Czech Republic also had a low inflation rate, low inflation volatility and a high household savings rate (Table 2). Inflation and the savings rate have important macroeconomic consequences on interest rates, including the rates charged on

mortgage loans. In the Czech Republic the real (after-inflation) interest rate on local currency deposits was negative between 2004 and 2008. This fact, together with the low inflation rate, allowed Czech mortgage lenders to offer low interest rates on local currency (CZK) mortgage loans. From 2005 to 2008, the average nominal rate on local currency loans was around 4-5% p.a.; for loans in the local currency for a five-year to ten-year fix period the average rate was around 5% p.a.

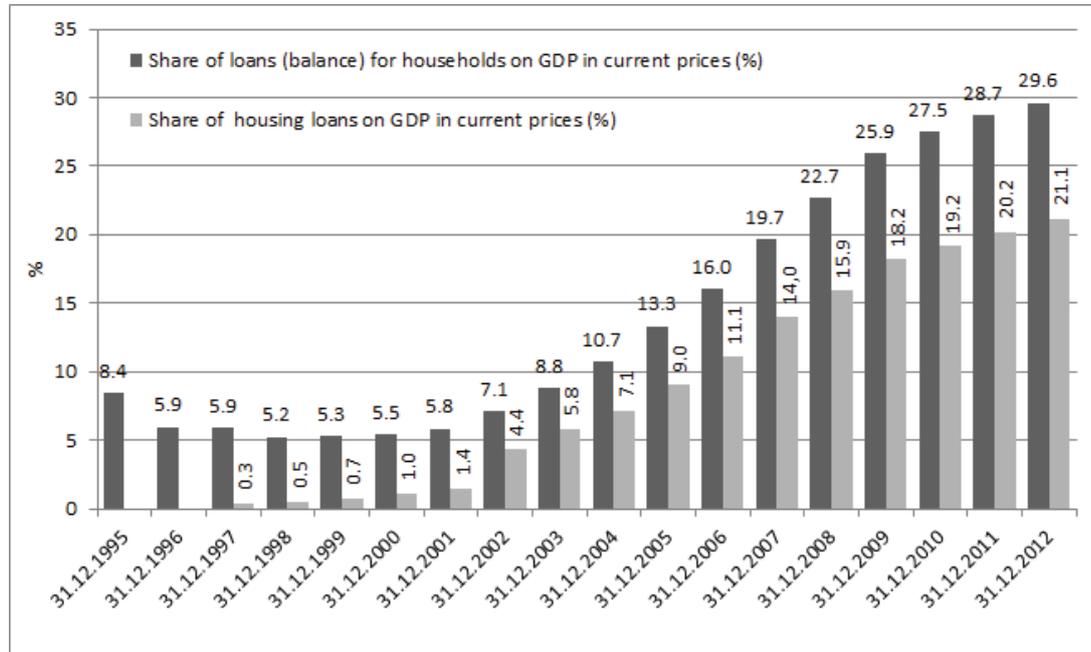
**Table 2: Contextual information, selected indicators (2005 – 2008)**

Socio-economic indicators	Years			
	2005	2006	2007	2008
<i>Economic indicators</i>				
GDP per capita in PPS (EU-27 average =100)	75.9	77.0	80.1	80.4
Growth of real GDP (% , y/y, real)	6.3	6.8	6.1	2.5
Household consumption (% , y/y, real)	2.5	5.0	4.8	3.6
Inflation rate (CPI, %, annual)	1.9	2.5	2.8	6.3
Unemployment rate (ILO, %, average)	7.9	7.1	5.3	4.4
Gross household saving rate	8.2	9.6	10.9	10.4
Total of housing savings (billions EUR)	11.1	12.7	13.9	16.1
Monetary base M0 (% , y/y)	11.4	11.9	10.0	12.8
M1 (% , y/y)	13.4	14.7	15.7	9.7
M3 (% , y/y)	11.1	13.8	16.7	13.1
General government deficit (surplus) / GDP (%)	-3.6	-2.6	-0.7	-2.7
General government debt as a % of GDP	29.7	29.4	29.0	30.0
<i>Income distribution, poverty rate</i>				
Income distribution: Gini	26.0	25.3	25.3	24.7
At-risk-of-poverty rate (%)	10.4	9.8	9.5	9.1
<i>Demography</i>				
Population (millions)	10.202	10.207	10.234	10.267
Proportion of population aged 25-49 years (%)	36.9	36.9	37.0	37.1
Proportion of population aged 65 and over (%)	14.2	14.4	14.6	14.9
Live births per 1,000 inhabitants	10.0	10.3	11.1	11.5
Natural growth per 1,000 inhabitants	-0.6	0.1	1.0	1.4
Total fertility rate	1.28	1.33	1.44	1.50

Note y/y denotes year-on-year change generally expressed in percentage points.

Source: Czech Statistical Office, Hungarian Central Statistical Office, Czech National Bank, National Bank of Hungary, Ministry of Finance of the Czech Republic, Eurostat

**Figure 3: Share of total loans and share of housing loans out of GDP in current prices (1995-2012)**



Source: Czech National Bank, Czech Statistical Office, own calculations.

Owing to low inflation combined with low local currency deposit/loan rates, the Czech mortgage loan portfolio contained a high share of fixed-rate mortgages (43% of all mortgages granted in 2008) and a marginal share of foreign currency loans (0.06% of all mortgages granted in 2008). The outstanding mortgage balance as a share of total GDP increased very steadily until 2001, but due to interest rate drops and deregulation of mortgage market (2004) it grew sharply between 2001 and 2008 and continued to grow until today (Figure 3).

Owing to high house price appreciation during the boom period (between 2000 and 2008 house prices increased by 123% according to the Czech Statistical Office price index) the attractiveness of the main substitute for owner-occupied housing, namely private rental housing with market rents, grew in the eyes of potential tenants. This is because market rents did not follow the same path of appreciation as house prices: the rent-to-price ratio in fact dropped substantially from 9 in 2000 to 5 in 2008. Private rental housing thus became more affordable and began to be characterized by more competition and a more broadly segmented supply (Lux and Sunega 2010).

In addition, the stock of public housing did not become marginalized: owner-occupied housing represented ‘only’ 61% of the housing stock (77% with housing cooperatives) and rental housing 23% of the housing stock in 2008 (13% private and 10.2% public rental housing).<sup>2</sup> In general, tenure choice was not as constrained as in other post-socialist states with ‘super-homeownership’ housing systems, and rental housing stock offered an adequate alternative of housing for those on lower or unstable incomes.

After Lehman Brothers filed for bankruptcy protection in the US on September 15, 2008, the ensuing uncertainty generated in the global financial markets and the simultaneous credit crunch had an immediate effect on financial markets in Central and Eastern Europe (CEE). In the Czech Republic, in late September and October 2008, there was a substantial drop in activity in the interbank money market, an increase in the volatility of exchange rates, with the Czech Crown (CZK) depreciating continuously as a result of reduced investor confidence in the whole CEE region until the second quarter of 2009, and a slump in the stock market. The Czech National Bank (CNB) reacted by lowering basic rates. For example, the rate for advances on collateral decreased cumulatively from August 2008 to December 2009 by 2.75 percentage points. However, these steps had only a limited effect on interbank money market rates and consequently on interest rates for mortgage loans. The monetary policy mechanism linking national bank rates to commercial rates was disrupted and risk premiums increased substantially.

Instead, banks tightened their loan conditions. In practice this meant that the loan financing of new housing development projects de facto stopped, and mortgage loans for households were tightened with the addition of a maximum loan-to-value ratio or by imposing a minimum income requirement on all new loan requests. The

---

<sup>2</sup> There are several factors that explain this more balanced tenure structure, such as: (1) the non-existence of a central right-to-buy policy (mass, quick and centrally directed privatization of public housing to the ownership of sitting tenants), and (2) strong public support for the restitution of housing stock in the Czech Republic during the early phases of the post-communist transition process. Support for public rental housing construction after 1990 is also of some relevance in this respect.

‘innovative’ mortgage products of the boom period (see Sunega and Lux 2007) quickly disappeared from the market, although formally banks continued to offer these products to customers. Owing to liquidity constraints there was a real danger of a run on Czech banks. This spurred the Czech government to increase state guarantees for deposits from EUR 25,000 to 50,000 in October 2008. In February 2010 the government doubled this state guarantee for deposits to EUR 100,000.

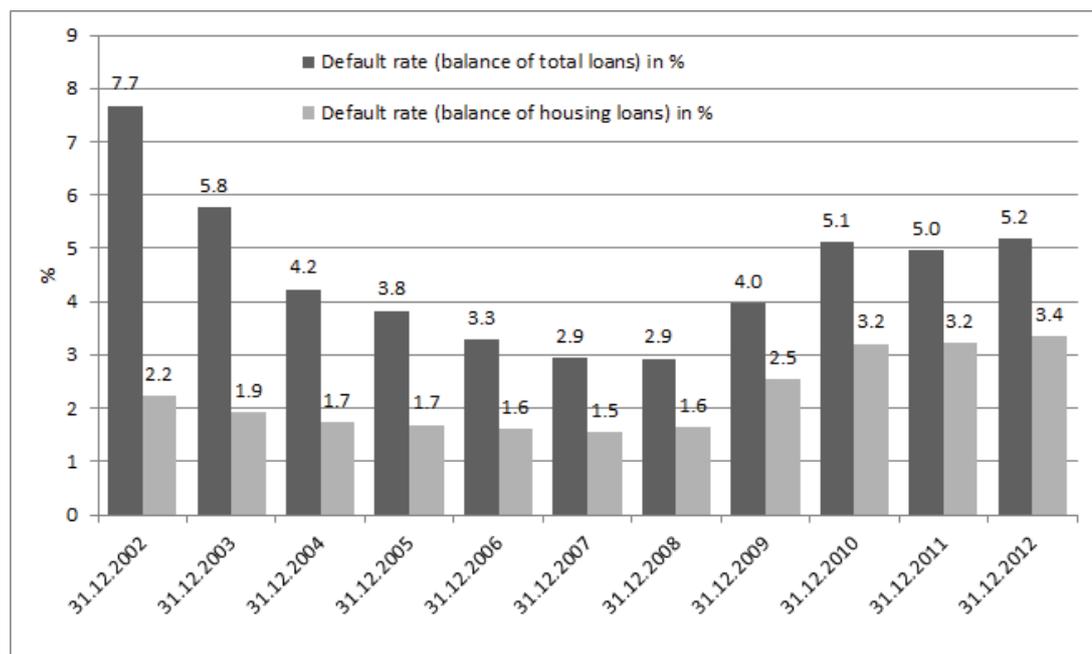
In 2008 the year-to-year GDP increase was only 2.5% (after 6-7% annual GDP growth between 2005 and 2007) and the economic recession, which started during the final quarter of 2008, led to an annual decline in GDP of 4.5% in 2009. In 2010, GDP increased by 3.1%, in 2011 by 1.8% but in 2012 it dropped again by 1%. However, the overall macroeconomic situation remained relatively stable in the Czech Republic. In 2009 annual inflation dropped to 1% (remained at the 1.5% level in 2010, 1.9% in 2011 and increased only in 2012 to 3.3%) and registered unemployment increased from 6% in 2008 to 9.2% in 2009 (and remained almost at the same level in 2010, slightly decreased to 8.6% in 2011 and increased to 9.4% in 2012).

In 2009 it became evident that the financial sector was in a healthy state, and despite the credit crunch Czech banks remained profitable. No Czech bank was taken over or went bankrupt during 2008-2010. It should be noted in this respect that the Czech Republic is one of the few new EU member states whose banking sector is independent of external financing, despite the fact that almost all the banks are owned by foreign capital (CNB 2009). The main reason for this situation is the high ratio of deposits to loans resulting from the comparatively high savings ratio of Czech households.

Figure 4 shows the trend in the default rate for all loans and housing loans (Bausparkasse and mortgage) separately. The default rate decreased between 2002 and 2008, increased in 2009 and 2010 and remained more or less stable in 2011 and 2012. During the global financial crisis and the economic recession in the Czech

Republic, the default rate on housing loans increased from 1.6% in 2008 to 2.5% in 2009 and to 3.4% in 2012. The acceptable increase was partly due to the relatively conservative lending conditions applied by Czech mortgage lenders until 2006, but more likely stemmed from the high share of fixed rate mortgages and the marginal share of foreign currency mortgages.

**Figure 4: The default rate for the balance of total loans and the default rate for the balance of housing loans (2002-2012)**



Source: Czech National Bank

Tighter loan conditions, decreasing household demand, falling exports, general economic recession and growing uncertainty in the labour market resulted in a drop in apartment prices (by 17% between 2008 and 2012 in nominal values according to the Česká spořitelna price index). However, prices of detached houses practically did not change and prices of plots increased by 10% in the same period.

The Czech government's response to the economic crisis in the financial sector, beyond the above-mentioned higher state guarantees for deposits, was limited. There were no mortgage rescue schemes, no special income support for highly leveraged borrowers, no new regulation of the banking sector, and nor moratorium on

repossessions either. The only government measures implemented were the postponement of rent deregulation in a number of larger Czech cities (the target date for full rent deregulation in 2010 was postponed to 2012), and the introduction of state guarantees for loans to housing developers building rental housing (however, there are very few rental housing development projects). Several employment policy measures such as tax incentives for employers with low-wage employees and measures designed to cut the public budget deficit were also introduced, but they did not directly or significantly influence the situation in the housing market.

## Conclusions

We described the background and the methodology for calculating the hedonic house price index based on data from Czech mortgage lenders, which is being gradually established since 2005. The price index built on data from mortgage lenders uses independent appraisals and supplement this database with information on region and location specific attributes (including commuting distances and times); it applies hedonic imputation method and both mechanic and statistical filtering to exclude outliers from the dataset. The price data are finally weighted for each combination of main characteristics; weights were derived from the census until 2013 and will be derived from transactions themselves since 2014. Since 2010, the dataset for price index is saturated by two leading Czech mortgage lenders and since 2013 by additional two smaller mortgage lenders, with about 60% total share on outstanding mortgage balance. The Česká spořitelna price index is suitable for measuring the current impact of the economic crisis on house prices. The impact of global economic crisis on mortgage and housing markets has been measured by trends in house prices and mortgage default rate. We showed that the impact of crisis on housing and mortgage markets in the Czech Republic was relatively mild up to now.

Our first hypothesis postulated that the impact the global economic crisis had on the mortgage and housing markets is explained by the macroeconomic environment and market regulations. This hypothesis, which is based on neoclassical economic theory, is confirmed by the empirical evidence presented in the preceding section. The Czech economy went through an economic crisis before in 1997-1998 and made significant fiscal and structural adjustments during the 2000 to 2002 period. In 2008, the Czech Republic had a high household savings ratio, low inflation and low inflation volatility. Czech mortgage lenders were consequently not dependent on external financing and could accumulate resources with relatively low deposit rates. The supply of credit consisted almost exclusively of local currency denominated mortgage loans and high share of fixed-rate mortgages; the interest rates for local currency mortgage loans remained low (between 4-5% p.a.).

Consequently, the economic factors that helped Czech mortgage market and banking system to remain in healthy conditions are following: (a) very short period of intensive mortgage boom before the crisis; (b) sufficient finance sources from deposits and mortgage bonds (also thanks to traditionally high household savings rate); (c) the rather low total indebtedness of Czech households; and (d) the fact that mortgage loans were denominated in local currencies and were mostly with fixed interest rates (thus well protecting borrowers from interest and exchange rate risks).

Our second hypothesis outlined above contends that the impact of the global economic crisis on the mortgage and housing markets is explained by wider institutional factors that influence the behaviour of market actors. Here emphasis is given to how the national housing system emerged and evolved. The evidence showed that in the Czech Republic the housing tenure structure remained balanced with relatively high shares of both public and private rental housing (while in many other post-socialist states public housing was marginalized and private rental housing now forms a small and high-priced sector). The yields (rent-to-price ratio) from residential rental investments dropped substantially and quickly, making rental housing a real alternative to owner-occupied housing. Housing subsidies in the Czech Republic, though also biased towards owner-occupied housing, have in the past also included support for public rental housing. This meant that new households on low or unstable income were not forced to become home-owners and had other tenure alternatives from which to choose permanent housing – the default rate thus remained relatively low even during the crisis period. From a comparison with Hungary (Hegedüs et al. 2011) it seems that this factor can explain an important part of the differences in the impact of the global economic crisis on the default rate and house prices between these two states.

In sum, the evidence presented above suggests that the second hypothesis can be accepted, with two caveats: the methodology used in this paper restricted the generality of the institutional mechanisms identified, and it was impossible to take

full account of other competing explanations. With these caveats in mind, there are still strong grounds to conclude that there are factors originating in the broader institutional frameworks (i.e. housing tenure policies and the development of housing systems) that explain the impact of the global economic crisis on the housing and mortgage markets in post-socialist states. In short, a full understanding of the impact of the global crisis requires a neoclassical explanation supplemented with an institutional account for otherwise puzzling features.

---

## References

- CNB, 2009. *Zpráva o finanční stabilitě* (Report on financial stability). Prague: Czech National Bank.
- Hegedüs, J., Lux, M., Sunega, P. 2011. Decline and Depression: The Impact of Global Economic Crisis on Housing Markets in Two Post-Socialist States, *Journal of Housing and the Built Environment* 26 (3): 315-333.
- Hodgson, G. 1998. The Approach of Institutional Economics, *Journal of Economic Literature* 36: 166-192.
- Lux, M., P. Sunega 2010. Private rental housing in the Czech Republic: Growth and...?. *Sociologický časopis/Czech Sociological Review* 46 (3): 349-373.
- North, D. 1991. Institutions, *Journal of Economic Perspectives* 5: 97–112.
- Sunega, P, Lux, M. 2007. Market-based housing finance efficiency in the Czech Republic, *European Journal of Housing Policy* 7: 241–273.



---

## 18. The hedonic housing price index for Poland – modelling on NBP BaRN data

Marta Wiślak



Marta Widłak – Economic Institute, Narodowy Bank Polski, ul. Świętokrzyska 11/21,  
00-919 Warszawa, Poland, marta.widlak@nbp.pl

The paper presents the personal opinion of the author and does not necessarily reflect the  
official position of the Narodowy Bank Polski.

---

## Article 18 Contents

1. Preface .....	207
2. Measurement of home price growth – adjustment for quality changes .....	209
3. Source of data .....	211
4. Index for Poland .....	213
5. Final conclusions and questions for further discussion .....	217
Literature .....	218
Appendix .....	219

## Abstract

The literature abounds with arguments regarding the relevance of reliable measurement of home prices. The correct measurement of home price growth is very important due to decisions taken by various economic agents and the impact of these decisions on the current situation in the domestic economy and its development.

The construction of a reliable home price index is a challenging task mainly due to the nature of the housing market and housing itself and availability of relevant data sources concerning this market. The most important issues which have to be addressed by a researcher involved in the measurement of home price trends, are mainly connected with the following two features of a good called housing and the housing market itself: heterogeneous nature of a housing good (there are no two identical housing goods) and rare turnover in housing stock (that is, relatively small number of transactions as compared to the housing stock volume).

This paper presents solutions to the abovementioned problems. Especially, the paper shows the construction of the hedonic house price index for Poland and describes in detail its modeling on NBP BaRN data.

## 1. Preface

The literature abounds with arguments regarding the importance of reliable measurement of home prices. In the past decade, a number of international research bodies were set up dealing with the problem of reliable measurement of home price growth.<sup>1</sup> Satisfactory arguments supporting the importance of reliable measurement of home price growth are given by Pollakowski and Meese and Wallace (1995, 1997 after: Prud'Homme et al. 2004). These include:

- home price growth is a key indicator of the situation in local markets;
- demographic decisions of the society, such as migration decisions and the choice of city and out-of-city residence are dependent on price trends in the housing market;
- residential properties are used as mortgage loan collateral; hence, home price trends are related with the valuation of risk borne by banks and other institutions financing the housing market ;
- potential homebuyers and investors are driven by home price growth in their search for property with price growth potential;
- makers of various kinds of economic policies based on price trends in the housing market, evaluate the merits of envisaged programmes such as housing schemes;
- correct measurement of real estate prices is necessary in the valuation of the aggregate assets in the economy;
- correct measurement of real estate prices is necessary to understand investment behavior in the economy and assess the effectiveness and cyclical nature of real estate markets.

To sum up, correct measurement of home price growth is very important due to decisions taken by economic agents and the impact of these decisions on the current situation in the domestic economy and its development. Such decisions are taken by:

- households,
- corporations and banks,

---

<sup>1</sup> Such bodies are set up within such organization as the OECD, BIS, IMF, Eurostat, central banks and statistical institutes.

- policy makers, including in particular: housing, monetary (including macro-prudential and financial stability), social and urban policies.

## 2. Measurement of home price growth - adjustment for quality changes

Construction of a reliable home price index is a challenging task mainly due to the nature of the housing units, housing market and availability of relevant data sources concerning this market. The most important issues which have to be addressed by a researcher analysing home price trends, are those connected with the following two features of a housing good and the housing market itself:

1. heterogeneous nature of a housing good (there are no two identical housing units)
2. low turnover rate in housing stock (that is, relatively small number of transactions as compared to the housing stock volume).

The above features determine two main issues associated with the construction of a reliable price index. The first issue concerns a **change (difference) in the quality of the analysed housing units**. The second one is representativeness of the index resulting from **representativeness of the data sample**.

Due to the heterogeneous nature of housing units and low turnover in housing stock, it is not possible to compare "like with like" in two different time periods. In particular, it is impossible to obtain the COLI-type index (cost of living index) - with fixed utility that sold dwelling units represent to their owners. Therefore, while constructing house price indices special attention should be paid to the index adjustment to reflect varying quality of goods sold (quality adjustment) resulting from heterogeneity of housing. **A reliable index of house prices should adjust for changes (differences) in the quality of housing units sold in different periods. Such an index will be considered to express the "pure" price change.**

The discussed issue of price index quality adjustment concerns various types of heterogeneous and rapidly changing goods. Therefore, it was widely studied in the context of the commonly used price indices such as the CPI or PPI. The affair of the so-called Boskin Commission<sup>2</sup> revealed that biases of price indices, caused, among other things, by non-

---

<sup>2</sup> The so-called Boskin Commission, officially referred to as "The Advisory Commission to Study the Consumer Price Index", was set up by the United States Senate in 1995 with the aim to study possible biases in the computation of the commonly used measure of inflation - the CPI. The main conclusions

adjustment for changes in the quality of goods can be considerable and can have serious economic consequences. The works of the Commission (Boskin et al. , 1996; 1998) show that the CPI in the United States was overstated in relation to the actual cost of living (COLI index) by approx. 1 % annually and in 1979 - 1995 by as much as approx. 1.3 % per year. The cumulative effect of the several years' revaluation would have, as argued by professors of the Boskin Commission, huge implications for the U.S. economic policy. Other interesting studies pointing to biases of price indices resulting from the substitution effect and changes in the quality of the analysed goods are cited by Hałka and Leszczyńska , 2010.

The problem of adjustment for quality changes in home price indices is usually solved, provided all relevant data are available, by the use of empirical hedonic models. A price index, which relies, in any way, on the information from the hedonic model is called a hedonic index (Triplett 2006). Another method to address changes in quality is to use the repeated sales index. This method may be applied in countries with very flexible labour and housing markets, where houses or apartments change their owners relatively quickly (e.g. the United States). Then, in order to investigate house price change, prices of resold housing are compared. This method is criticized mainly for failing to take account of changes in the quality of resold residential properties (refurbishment and improvements performed in the meantime) and for the sample selection bias (properties of inferior quality or specifically located properties are usually sold quickly).

---

of the Commission's work were presented in a report entitled "Toward A More Accurate Measure Of The Cost Of Living" (see Boskin et al., 1996).

### 3. Source of data

The issue of appropriate measurement of home price growth is inextricably linked to the availability of relevant data. Despite its simplicity, the concept of hedonic adjustment for quality changes cannot often be implemented due to the lack of relevant data bases containing both, information on house prices and a description of the most important house characteristics. These should include a full address to enable accurate house geocoding and analysis of spatial relationships (use of methods of spatial econometrics or extension of the list of price factors to include specific characteristics of a particular location).

For the purpose of signal- analysis and monitoring of price trends across countries, in practice, very different sources of information about residential real estate are used. They differ not only in the scope of house description (characteristics) but also in the frequency of accumulation of observations and sources of original information. Typically, the collected information can be divided according to their source of origin (Łaszek , Widłak 2008) as follows:

- notarial deeds,
- real estate agencies,
- tax statistics,
- database associated with common taxation of real estate,
- data collected by the banking system (valuation and transaction data),
- appraisals and individual research.

These sources have their advantages and disadvantages. Considering the theoretical hedonic model that is basis for quality adjustment in price indexes, transaction data are the most appropriate one. Only those data are the result of preferences and decisions of both parties to the transaction - the seller and the buyer. Unfortunately, in most countries problems associated with the analysis of transaction prices result from price understatement in notarial deeds, done for tax evasion purposes, and delays associated with legal and administrative procedure of home purchase and signing the final notarial deed (notarial deed is usually signed much later than the original agreement on prices is reached - and therefore the price stated in the notarial deed reflects current market prices with a time lag).

Very often due to problems associated with transaction data, asking price data are used. In such a situation it should be remembered that unlike transaction prices, they reflect only one side of the market – preferences of the seller. It seems that in the model terms, the asking price is determined solely by the utility function of the seller. This means that if the transaction price of housing is actually the result of supply and demand force, the asking price reflects the supply side of the market only. The microeconomic hedonic model of price formation in the housing market (see para . 3.2 ) assumes that the price is the result of supply and demand for a particular good. Therefore, in order to analyse the actual price trends in the housing market it is best to use transaction prices.

In the situation of imbalance in the market, namely demand exceeding supply, asking price growth may be a good approximation of trends in transaction prices. During periods of strong price increases caused by excessive demand as compared with supply, asking price levels actually get closer to transaction prices. For example, in Poland this phenomenon was observed in 2005-2007 in the seven largest cities (see NBP 2010, 2011, 2012). In the situation where the market is close to equilibrium, as a rule, the level of asking prices is higher than the level of transaction prices; sellers take into account buyer's readiness to negotiate and deliberately inflate the asking price above the expected transaction price. The asking price index can serve as a leading indicator in the assessment of housing market situation and, in the absence of transactional data, can be used to evaluate market price trends.

Appraisals, particularly those made in the banking sector, intended to estimate the value of collaterals of mortgage loan portfolio are becoming an increasingly common source of information in home price research. Yet, this source does not seem appropriate as it does not reflect true market information. Property valuation is always connected with the subjective role of the property appraiser, and hedonic models based on data from this kind of source reflect weights attributed to individual characteristics by property appraisers rather than implicit prices of property characteristics determined by real supply and demand forces.

#### 4. Index for Poland

Due to the lack of relevant data, Poland's experience in computing the home price index is very limited. Since 2006, Narodowy Bank Polski has studied housing prices, among other things, in order to obtain reliable measurement of price trends. Until now, the studies have focused on choosing an appropriate index construction method and on analyzing different hedonic methods for quality adjustment. Below, we have presented the research results for the Warsaw market and 16 voivodship cities. The findings for Poland's largest housing market, Warsaw, help us to choose the hedonic method for quality adjustment. At the beginning four different methods of computing hedonic price indices were presented: imputation method (M4), characteristics price method (M3) and methods based on time dummy variables of adjacent periods (M2) and pooled (all periods) approach (M1).<sup>3</sup> Then, using the method of time dummy variables (M1) we estimated the aggregate index for other 15 cities – voivodship capitals, where NBP monitoring of real estate prices is conducted.

The following analysis relies on the data collected by NBP, which monitors both asking and transaction prices in the housing markets of Poland's largest cities. The presented price index concerns transactions in the secondary housing market.

Tables 2 and 4 show the results of selected hedonic models, which were used to determine the indices. The models tested, at the same time, the impact of the so-called spatial variables related to the characteristics of the neighborhood and the wider environs of the property. Out of 28 different spatial indicators for Warsaw, finally 13 indicators in model M1 and 6 in model M4 with a bigger number of degrees of freedom, were considered important and having an economically reasonable value. It should be remembered that the estimates for these variables are subject to an error, since the geocoding of the properties was done with an accuracy to the street midpoint (the NBP data base does not contain accurate address data of the properties; the study included streets which are maximum 3 kilometers long). Among numerous estimated models, it was concluded that the models taking into account both the location in a particular city district and characteristics of the immediate neighborhood as well as the general location on the city map (models marked in red in Table 2) are

---

<sup>3</sup> Detailed methodology for determining the indices can be found in Widłak, 2010.

the most suitable to the real data. As suggested during the conference, the chosen third version of M1 model (v3, see Table 2) may be "too sophisticated" due to a large number of variables describing the space characteristics. Experience in hedonic modeling shows that the property location in the city district, its distance from the city centre (city limit) and the underground are, apart from structural features of the property, the most important price factors. An interesting observation is that, according to the analysis of the real data, the location of the district is important per se, and parameters such as the distance from the city center, underground stations, shopping malls, parks, etc. are factors affecting the price regardless of the property location in the district (Previously it was assumed that the location in the district determines at least some of these correlations, such as distance to the center and underground stations). Descriptive statistics of the key variables used in the modeling are shown in Tables 1 and 3 of the Annex.

Figure 1 shows average levels of home prices in Warsaw from 2006 to 2013 as measured with the use of three different methods - simple average, weighted average (with fixed weights for particular districts) and the average prices of housing units sold in 2006 Q1 adjusted with the hedonic index. Figure 1 additionally includes an index of the structure of housing characteristics.

This indicator tells us what would be the cost of changing (from quarter to quarter) housing sample (changing housing characteristics structure) if implicit prices of housing characteristics are constant (determined as average prices of housing within a given category of characteristics in 2011). This indicator refers to all the data and is based on the same data sample as taken for the simple and weighted average. The comparison of the presented indicators suggests the following conclusions. Firstly, changes in the average price measure and the structure index are very similar. This suggests a significant correlation of the average home price index with the quality of analysed housing. In 2010 Q1, we observed a marked structural change in the distribution of characteristics involving a permanent shift towards "inferior" quality housing. (Starting with this quarter, the database contains relatively more transactions involving housing units situated in the district of Targówek, built in 80s. in prefabricated large panel construction technology and with lower standard of finishing - see Figures 3 - 4). This change in the structure of sold housing does not concern the entire, unknown population of transactions (samples in the BaRN database are not randomly chosen).

This change is then reflected in the decline of the level of average prices in Warsaw. On the other hand, the average level of prices of housing sold in 2006 Q1 and adjusted with the hedonic index (the so-called hedonic price) is neither subject to such significant fluctuations as average price levels nor to change of characteristics structure. This comparison also shows, and this is the second important conclusion, that the hedonic index indeed adjusts for differences in the quality of sold housing as its variability from quarter to quarter is not as large as that of average prices. Even if we calculate average prices based on the same sample, which we use to construct the hedonic index (no structural shift in the quality of housing), we can still see an evident correlation between the average home price growth and the change in the structure index (see Figure 2). Aside from the first three quarters, the average price growth changed in line with the changes in the structure of housing characteristics. This correlation is more pronounced in the case of the simple average than the hedonic index. This latter also has, as noted above, a smaller variance.

On the other hand, the comparison of hedonic indices constructed using different methods (see Figure 5) suggests the following conclusions. The characteristics price method, theoretically considered to be the most accurate one (see Widłak 2010) is not an appropriate measure to study price trends in the Warsaw market. As shown in the Figure, this index is subject to huge fluctuations, not supported by other indicators. Implicit prices of housing characteristics in quarterly models change significantly, which is reflected in the index variance. Implicit prices of characteristics change substantially from quarter to quarter but these changes do not coincide with changes in structure of characteristics indices (see figures 3-4 and 9 - 40). It seems that such significant changes, suggesting fast and vast differences in consumer tastes are not fully justified. Partially, market imperfection and variability of tastes, and partially, difficulties with correct specification of these models (a lot of regressions' coefficients are statistically insignificant at the 10% level) and data errors may account for the problem.

Therefore, in the case of the analysed markets, it is better to use the time dummy variable method or imputation method. As shown in Figure 6, indices constructed with the use of these methods are the least variable, which suggests that they are better adapted to changes in the quality of dwellings sold. At the same time, the method combining data from all

periods (M1) is subject to less fluctuation than adjacent period approach (M2), and the index revisions resulting from the definition of this method seem to be negligible (see Figure 7).

Price growth for the total group of 16 voivodship cities was determined with the use of the pooled time dummy method (M1). The cities were divided into a group of 6 cities (Warsaw, Kraków, Tricity (Gdańsk-Gdynia-Sopot metropolitan area), Wrocław, Poznań, Łódź) and 10 cities (Białystok, Bydgoszcz, Katowice, Kielce, Lublin, Olsztyn, Opole, Szczecin, Rzeszów, Zielona Góra). The aggregate price index is represented by an average weighted with the share of both groups of cities in the total of transactions collected in the BaRN database (see Figure 8). All the indicators reflect the situation in the housing market in the recent years. The Figure suggests that there is no considerable delay in price growth in smaller cities as compared to the trends observed in the largest cities. Price trends are similar for both groups of cities, with markedly different price levels, which also affects the level of price indices in both groups and the visible base effect.

## 5. Final conclusions and questions for further discussion

The most important general conclusions from the simulations presented in the article are as follows.

- The analysis of the structure of housing characteristics and price indices (simple average, weighted average with fixed weights for city districts and the hedonic index) implies that the hedonic approach is the most reliable one. The hedonic index is the best way to adjust for changes in the housing characteristics structure that occur from quarter to quarter (*housing heterogeneity*).
- The need to address the data selection bias (randomization of samples needed but difficult to plan as the entire population of transactions is unknown – lack of official, reliable data).
- One of the questions raised in the analysis was which hedonic approach should be applied. The detailed analysis of the Warsaw market shows that imputation and pooled time – dummy indices are preferable (the imputation method is not the same as the characteristics price method as described in Widłak 2010). However, an answer to the question why these two approaches give similar results despite different theoretical assumptions requires further investigation. Unfortunately, not much experience is shared internationally in this field. Another problematic question resulting from this survey is why implicit prices estimated for the characteristics price method are so unstable and make the index so volatile. The above excludes this theoretically best approach from empirical use.

All in all, the pooled-time dummy approach was the preferred method to calculate the composite index (for 6 and 10 voivodeship capitals). This is also due to operational reasons (no need for re-specification and re-estimation of models on a quarterly basis; revisions are not so important, either).

## Literature:

1. Information on home prices and the situation in the residential and commercial real estate market in Poland – quarterly reports (2010 - 2013), NBP,  
[http://www.nbp.pl/homen.aspx?f=/en/publikacje/inne/real\\_estate\\_market\\_q.html](http://www.nbp.pl/homen.aspx?f=/en/publikacje/inne/real_estate_market_q.html)
2. Widłak M. (2010), *Metody wyznaczania hedonicznych indeksów cen jako sposób kontroli zmian jakości dóbr*, Wiadomości Statystyczne, No. 9 (592)
3. Łaszek J., Widłak M. (2008), *Badanie cen na rynku mieszkań prywatnych zamieszkałych przez właściciela z perspektywy banku centralnego*, Bank i kredyt.
4. Boskin M. J., Dulberger E. R., Gordon R. J., Griliches Z., Jorgenson D. W. (1998), *Consumer Prices, the Consumer Price Index, and the Cost of Living*, Journal of Economic Perspectives, Vol. 12, No. 1.
5. Hałka A., Leszczyńska A. (2011), *Wady i zalety wskaźnika cen towarów i usług konsumpcyjnych – szacunki obciążenia dla Polski*, Gospodarka Narodowa, No. 9/2011.
6. Prud'Homme, Sanga D., Shum H. (2004), *From average to hedonic price indexes: a "preliminary" investigation into various measures of trends in existing house prices using MLS Data for Ottawa*, paper prepared for 28<sup>th</sup> General Conference of the International Association for Research in Income and Wealth, Cork, Ireland.

## Appendix

Source of all tables and figures is own calculations on NBP BaRN data.

Table 1: Descriptive statistics for the Warsaw data.

Variable	mean	p50	min	max	sd
Price per sq. m.	7689	7499	2800	16904	1798
Surface (sq m.)	53	48	14	229	24
Sq_surface (sq m.)	3413	2350	196	52441	3755
Rooms 1 or 2	0,64	1	0	1	0,48
High finishing standard	0,33	0	0	1	0,47
Low finishing standard	0,23	0	0	1	0,42
Cooperative ownership	0,42	0	0	1	0,49
Construction year <1945	0,06	0	0	1	0,24
Construction year 1945 - 1970	0,30	0	0	1	0,46
Construction year 1970 - 1979	0,17	0	0	1	0,38
Construction year 1980 - 1988	0,10	0	0	1	0,29
Construction year 2002 - 2005	0,11	0	0	1	0,31
Construction year > 2005	0,10	0	0	1	0,29
Under construction	0,03	0	0	1	0,18
Białołęka, Targówek	0,10	0	0	1	0,30
Praga-Północ, Praga-Południe	0,10	0	0	1	0,30
Rembertów, Wesoła, Wawer	0,01	0	0	1	0,10
Włochy, Ursus	0,04	0	0	1	0,19
Mokotów	0,16	0	0	1	0,37
Ochota	0,04	0	0	1	0,21
Wola	0,09	0	0	1	0,29
Ursynów	0,19	0	0	1	0,39
Żoliborz	0,03	0	0	1	0,18
Wilanów	0,01	0	0	1	0,12
Bielany, Bemowo	0,10	0	0	1	0,30
Log city boundry dist.	4484	4537	57	8397	2052
Log metro station dist.	2249	1436	22	10569	2161
Tram noise	13,56	0	0	70	22,36
Industrial noise	3,15	0	0	55	11,06
Green areas in 1000 m. circle	16,59	15	1	63	8,98
Cemetery in 500 m. circle	0,13	0	0	3	0,43
Discounter in 1000 m. circle	0,71	0	0	5	0,90
Medical center in 1000 m. circle	5,10	5	0	24	3,92
Hospitals in 1000 m. circle	3,56	1	0	29	5,66
Log green area dist.	251	222	3	906	169
Log protected green points dist.	3360	2769	16	10000	2298
Log industrial points dist.	602	510	13	2448	433
Log trade center dist.	1331	1070	12	10602	1178

Table 2: Hedonic models for Warsaw.

Variables	Pooled time - dummy (v1 M1)	Pooled time - dummy (v2 M1)	Pooled time - dummy (v3 M1)	Imputation (v1 M4 2011)	Imputation (v2 M4 2011)	Imputation (v3 M4 2011)
Surface (sq m.)	-0,005 ***	-0,004 ***	-0,005 ***	-0,006 ***	-0,006 ***	-0,007 ***
Sq_surface (sq m.)	0 ***	0 ***	0 ***	0 ***	0 ***	0 ***
Rooms 1 or 2	0,011 *	0,021 ***	0,014 **	0,039 ***	0,02	0,002 ***
High finishing standard	0,065 ***	0,062 ***	0,06 ***	0,061 ***	0,052 **	0,054
Low finishing standard	-0,048 ***	-0,056 ***	-0,051 ***	-0,077 ***	-0,067 ***	-0,067 ***
Cooperative ownership	-0,016 ***	-0,013 ***	-0,018 ***	-0,024 **	-0,005	-0,015 ***
Construction year <1945	0,008 **	-0,028 **	-0,041 ***	0,094 ***	-0,05 *	-0,082 **
Construction year 1945 - 1970	-0,083 ***	-0,109 ***	-0,127 ***	-0,06 ***	-0,122 ***	-0,141 ***
Construction year 1970 - 1979	-0,13 ***	-0,11 ***	-0,143 ***	-0,081 ***	-0,113 ***	-0,144 ***
Construction year 1980 - 1988	-0,14 ***	-0,11 ***	-0,143 ***	-0,073 ***	-0,1 ***	-0,135 ***
Construction year 2002 - 2005	0,069 ***	0,055 ***	0,064 ***	0,055 ***	0,036 **	0,042 ***
Construction year > 2005	0,109 ***	0,105 ***	0,11 ***	0,143 ***	0,094 ***	0,088 ***
Under construction	-0,03 **	-0,005	-0,02	-0,081	-0,022	-0,021 ***
Bielany, Bemowo	-0,262 ***		-0,144 ***	-0,264 ***		-0,157
Białołęka, Targówek	-0,444 ***		-0,325 ***	-0,416 ***		-0,342 ***
Praga-Północ, Praga-Południe	-0,292 ***		-0,194 ***	-0,258 ***		-0,21 ***
Rembertów, Wesoła, Wawer	-0,502 ***		-0,367 ***	-0,524 ***		-0,426 ***
Włochy, Ursus	-0,357 ***		-0,18 ***	-0,341 ***		-0,173 ***
Mokotów	-0,123 ***		-0,094 ***	-0,135 ***		-0,113 ***
Ochota	-0,153 ***		-0,114 ***	-0,138 ***		-0,094 ***
Wola	-0,195 ***		-0,108 ***	-0,214 ***		-0,118 ***
Ursynów	-0,153 ***		-0,085 ***	-0,172 ***		-0,129 ***
Żoliborz	-0,13 ***		-0,096 ***	-0,137 ***		-0,035 ***
Wilanów	-0,236 ***		-0,155 ***	-0,229 ***		-0,204
Log centrum dist.		-0,112 ***			-0,09 **	
Log city boundry dist.		0,034 ***	0,052 ***		0,031 **	0,037 ***
Log metro station dist.		-0,055 ***	-0,025 ***		-0,054 **	-0,024 ***

Tram noise		-0,001 ***	001 ***		-0,001 ***	
Industrial noise		-0,001 ***	-0,001 ***		-0,001 ***	-0,001 ***
Green areas in 1000 m. circle		0,002 ***	0,002 ***		0,002 ***	0,001 **
Cemetery in 500 m. circle		-0,008 *	-0,022 ***			
Discounter in 1000 m. circle		-0,004 *	0,005 **			
Medical center in 1000 m. circle		0,003 ***	0,005 ***			
Hospitals in 1000 m. circle		0,001	0,002 ***			
Medical centers + hospitals in 1000 m. circle					0,003 ***	0,004 ***
Log green area dist.		-0,007 ***	-0,012 ***			
Log protected green points dist.		0,005	0,033 ***			
Log industrial points dist.		0,011 ***	0,007 ***			
Log trade center dist.		0,009 ***	0,010 ***		0,009 *	0,011 **
Const.	8,871 ***	9,548 ***	8,202 ***	9,333 ***	10,018 ***	9,195 ***
Number of obs.	4955	4073	4070	824	862	862
R2	0,6899	0,6886	0,7293	0,6056	0,5699	0,6614
R2 (adj)	0,6865	0,6842	<b>0,7248</b>	0,5937	0,5591	<b>0,6488</b>
AIC	-6059	-4870	<b>-5444</b>	-1046	-1228	<b>-1421</b>
Ramsey's RESET	0,0832	0,0342	<b>0,2418</b>	0,1383	0,0232	<b>0,0412</b>

Note: Data for coefficients of time dummies in M1 approaches are not quoted in the table (they could be seen at the relevant figures below).

\* - stands for statistical significance of a coeff. at 10%, \*\* - 5%, \*\*\* - 1%.

Table 3: Geographical distribution of the sample.

City	No of obs.	percent
Warszawa	8 389	12,1
Wrocław	7 203	10,4
Kraków	6 615	9,5
Szczecin	6 188	8,9
Olsztyn	5 395	7,8
Poznań	4 827	7,0
Gdańsk	3 713	5,4
Lublin	3 393	4,9
Opole	3 139	4,5
Białystok	3 056	4,4
Kielce	2 982	4,3
Bydgoszcz	2 912	4,2

Table 4: Hedonic pooled time dummy models (M1) for the group of 6 and 10 cities.

Variables	Model M1 for 6 cities (Warsaw is ref.)	Model M1 for 10 cities (Szczecin is ref.)	
An average location in a neighbouring area	0,104 ***	An average location in a neighbouring area	0,088 ***
Good location in a neighbouring area	0,168 ***	Good location in a neighbouring area	0,131 ***
High finishing standard	0,095 ***	High finishing standard	0,101 ***
-	-	-	-
Low finishing standard	0,096 ***	Low finishing standard	0,119 ***
-	-	-	-
Tradditional construction technology	0,036 ***	Tradditional construction technology	0,013 **
-	-	-	-
Monolit construction	0,089 ***	Monolit construction	0,029 ***
-	-	-	-
Prefabricated construction	0,055 ***	Prefabricated construction	0,048 ***
-	-	-	-
1st floor	0,027 ***	1st floor	0,023 ***
-	-	-	-
last floor	0,013 ***	last floor	0,018 ***
-	-	-	-
No of floors	0,000	No of floors	0,004 ***
-	-	-	-
Kitchen with living room	0,045 ***	Kitchen with living room	0,027 ***
-	-	-	-
Cooperative ownership	0,017 ***	Cooperative ownership	0,007 ***
-	-	-	-
Construction year <1945	0,072 ***	Construction year <1945	0,157 ***
-	-	-	-
Construction year 1945 - 1970	0,028 ***	Construction year 1945 - 1970	0,079 ***
Construction year 1970 - 1979	- ***	Construction year 1970 - 1979	- ***

	0,037		0,056
	-		-
Construction year 1980 - 1988	0,037 ***	Construction year 1980 - 1988	0,061 ***
Construction year 2002 - 2005	0,057 ***	Construction year 2002 - 2005	0,064 ***
Construction year > 2005	0,108 ***	Construction year > 2005	0,071 ***
Rooms1	0,121 ***	Sq_surface	0,000 ***
Rooms 2		Surface	-
	0,074 ***		0,005 ***
Rooms 3		Bialystok	-
	0,025 ***		0,050 ***
Gdansk	-	Bydgoszcz	-
	0,366 ***		0,157 ***
Gdynia	-	Katowice	-
	0,381 ***		0,169 ***
Sopot		Kielce	-
	0,040 ***		0,069 ***
Krakow	-	Lublin	
	0,261 ***		0,061 ***
Poznan	-	Olsztyn	-
	0,361 ***		0,005
Wroclaw	-	Opole	-
	0,308 ***		0,084 ***
Lodz	-	Rzeszow	-
	0,685 ***		0,006
		Zielona_Gora	-
			0,326 ***
Number of obs.	10583		14861
R2	0,6964		0,6094

Note: The same notes as for table 2.

Figure 1 Average home price level and the structure index of housing characteristics (Warsaw, RWT).

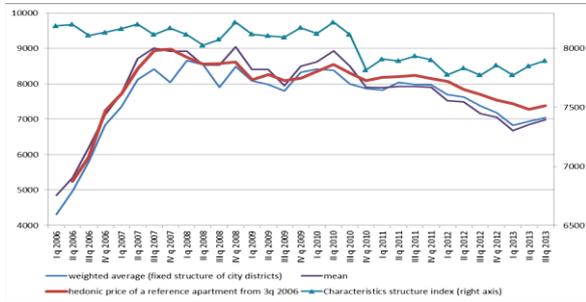


Figure 2 Home price growth and the structure index of housing characteristics (previous quarter = 100).

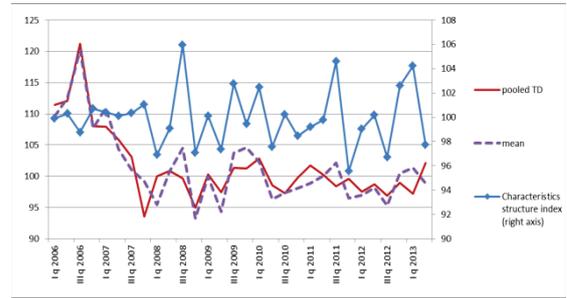


Figure 3 Indices of the structure of selected housing characteristics (2011 constant prices) Part 1.

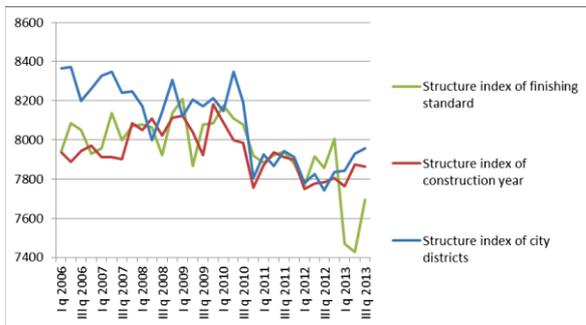


Figure 4 Indices of the structure of selected housing characteristics (2011 constant prices) Part 2.

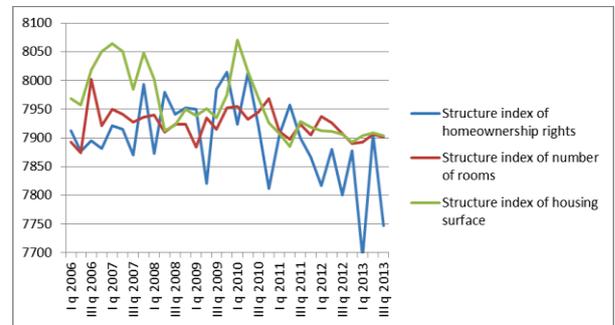


Figure 5 House price growth – hedonic indices and simple average index (Warsaw, RWT, previous quarter = 100).

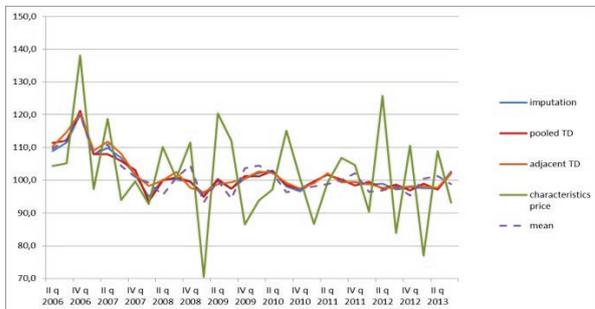


Figure 6 House price growth – hedonic indices (exclusive of characteristics price index) and simple average index (Warsaw, RWT, previous quarter = 100).

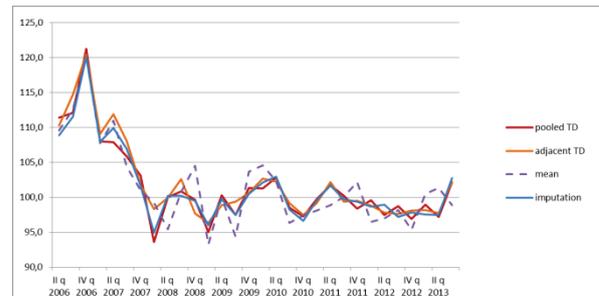


Figure 7 Value of pooled time dummy index (M1) (2006 Q1 = 100) determined in subsequent periods (reflects possible revisions).

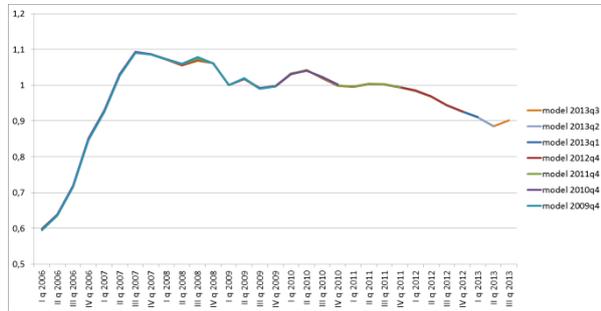


Figure 8 House price growth in 16 voivodship cities (2006 Q3 = 100).

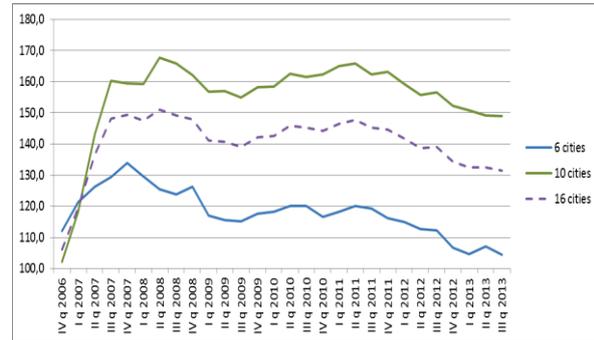
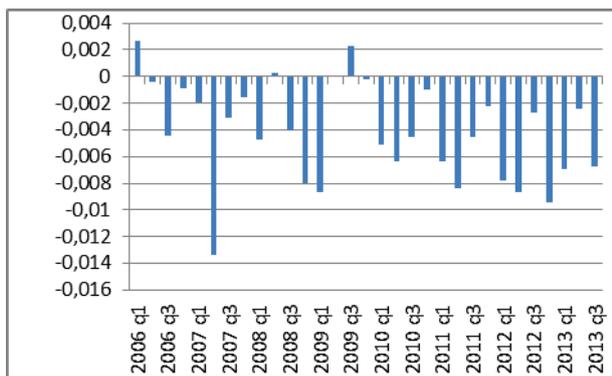
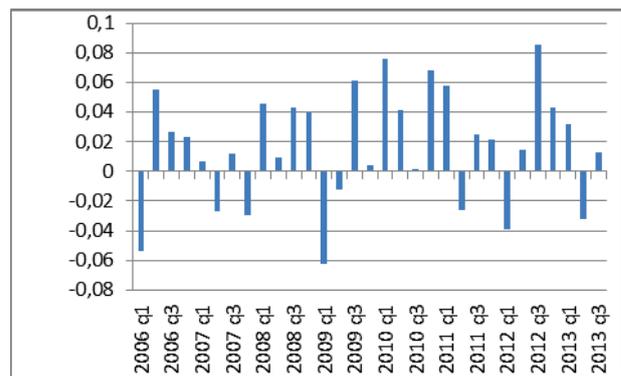


Figure 9 – 40: Implicit prices of the house characteristics estimated for characteristics' price approach (M3 models):

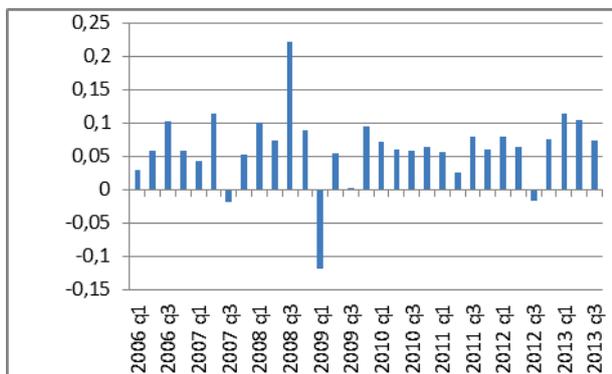
Surface



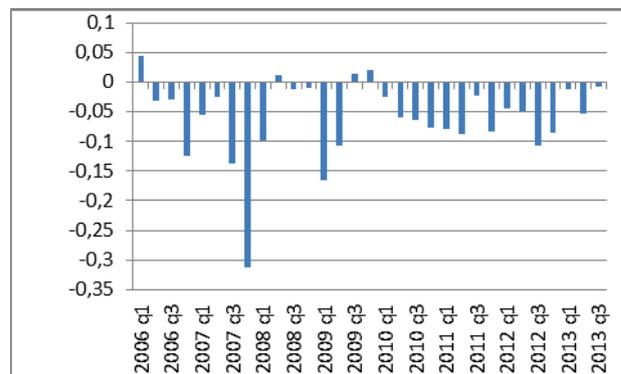
Rooms 1 or 2



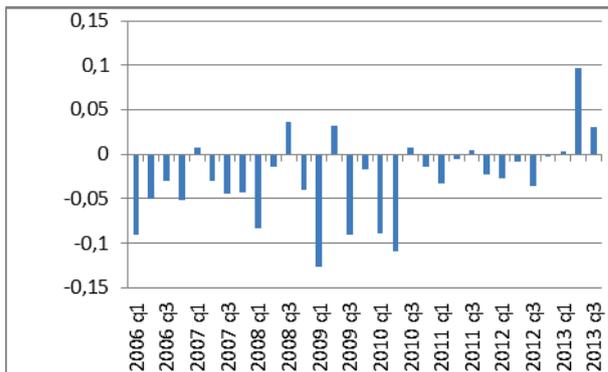
High finishing standard



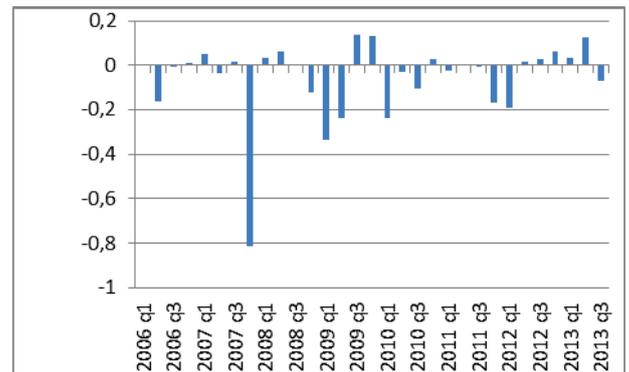
Low finishing standard



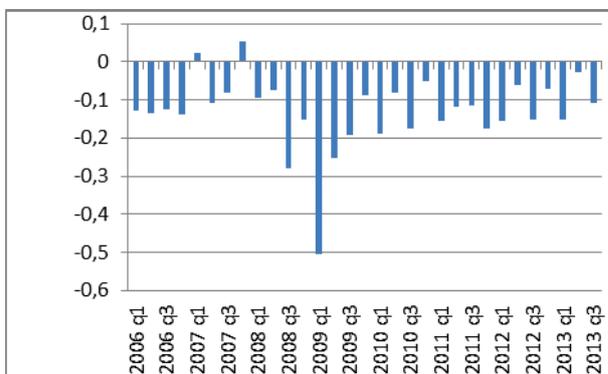
Cooperative ownership



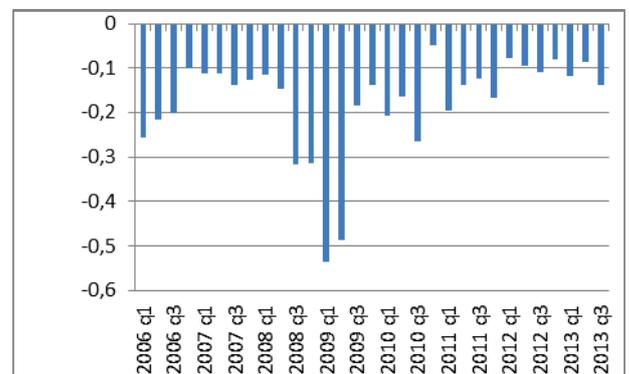
Construction year <1945



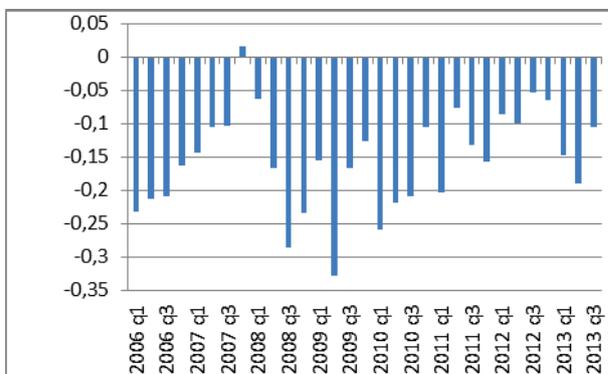
Construction year 1945 - 1970



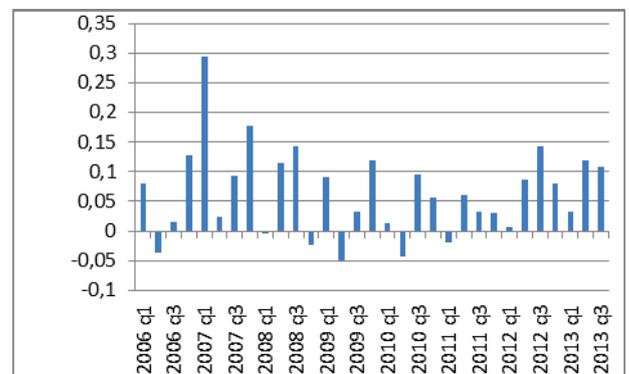
Construction year 1970 - 1979



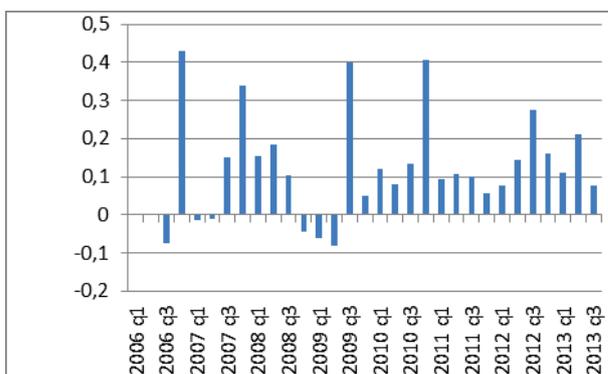
Construction year 1980 - 1988



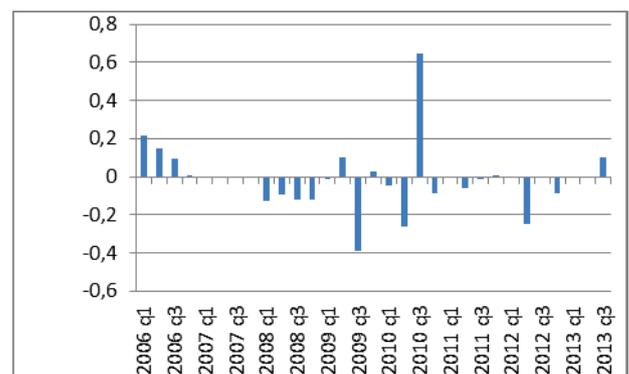
Construction year 2002 - 2005



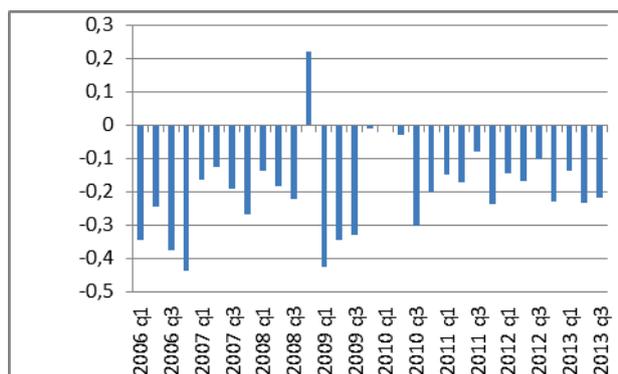
Construction year > 2005



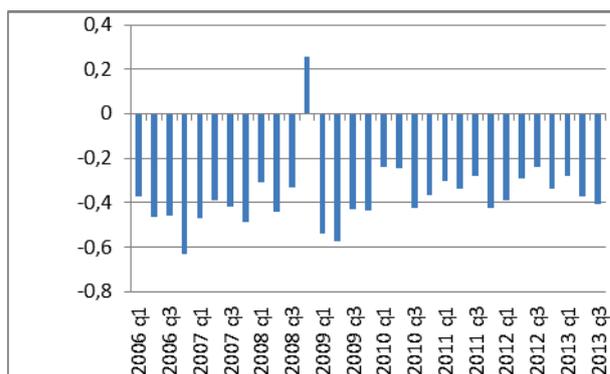
Under construction



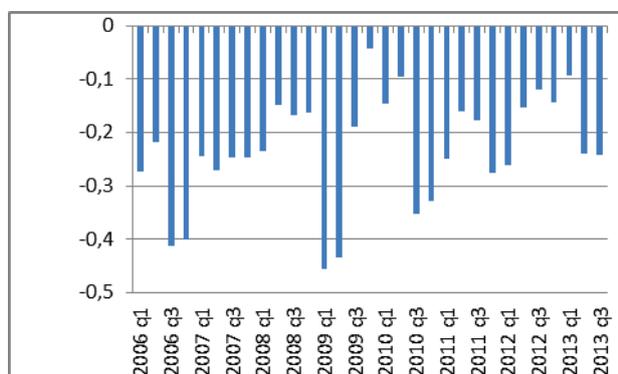
Białoteka, Targówek



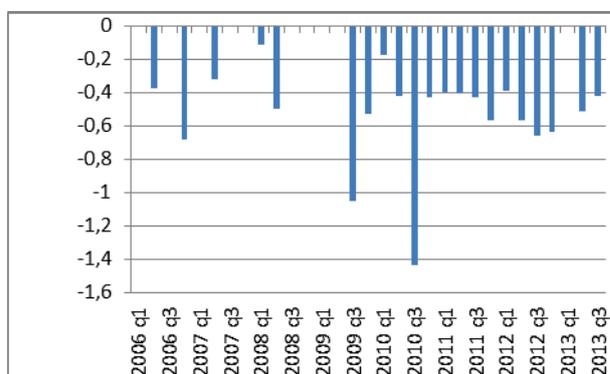
Praga-Północ, Praga-Południe



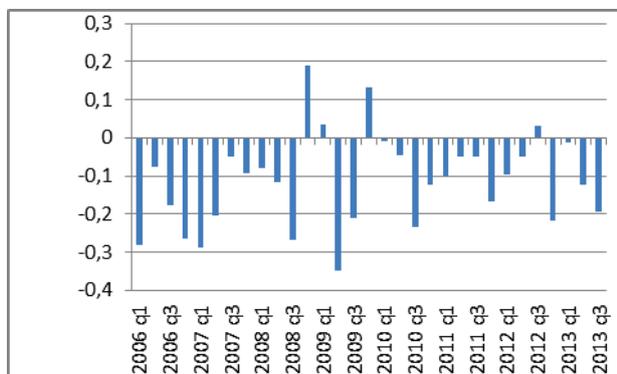
Rembertów, Wesola, Wawer



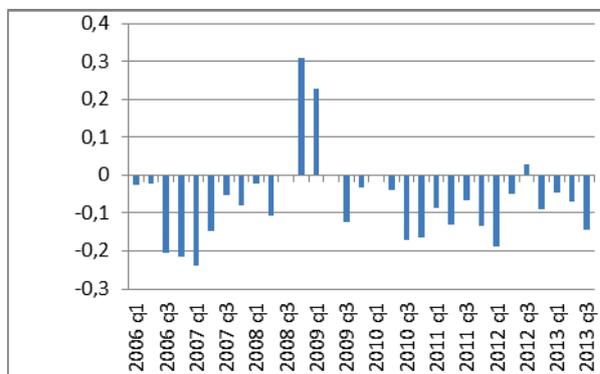
Włochy, Ursus



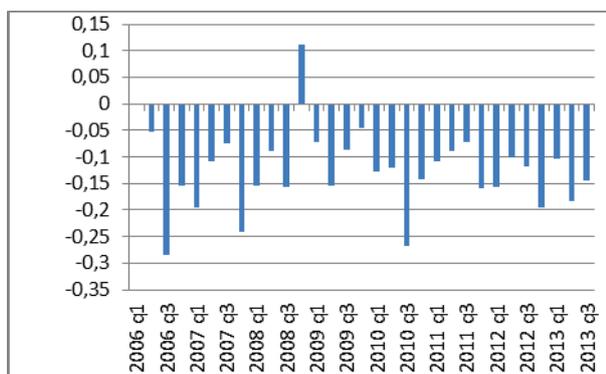
Mokotów



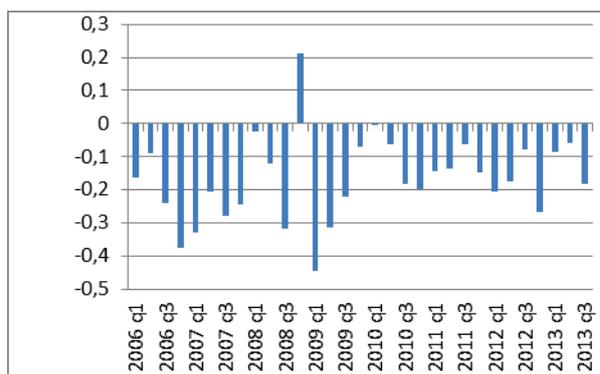
Ochota



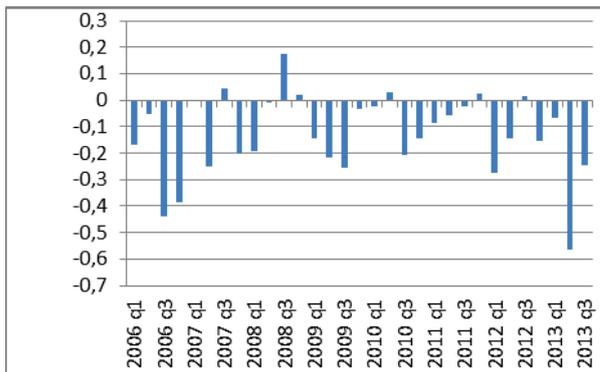
Wola



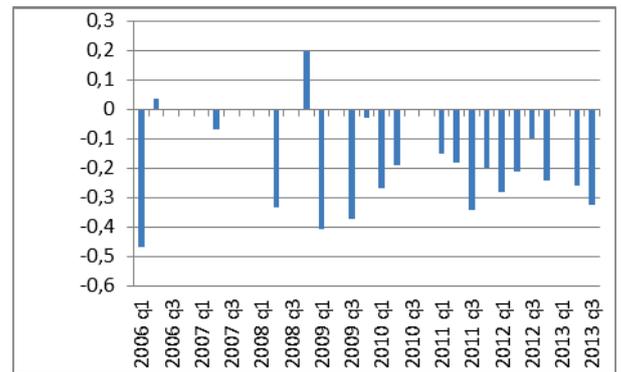
Ursynów



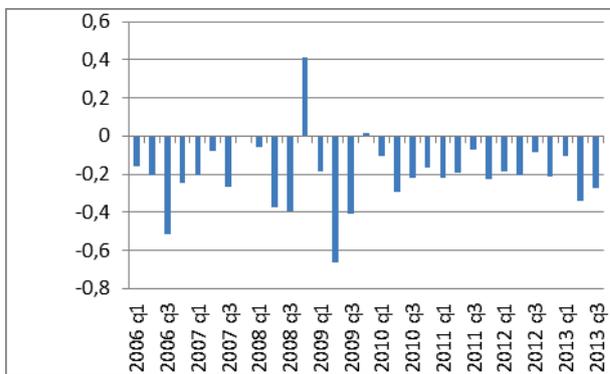
Żoliborz



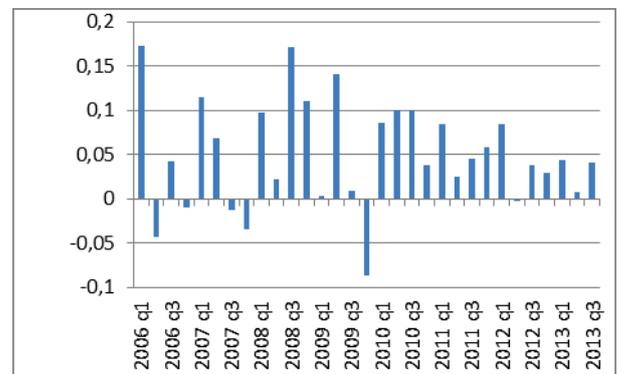
Wilanów



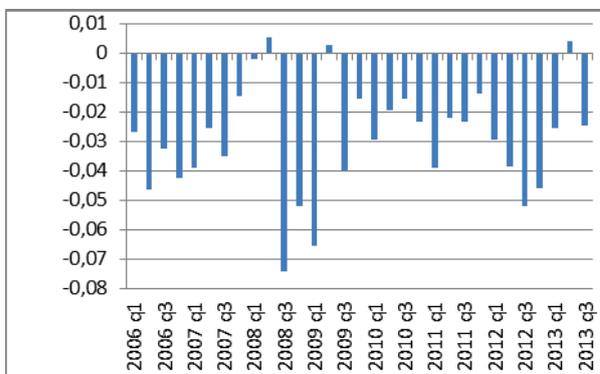
Bielany, Bemowo



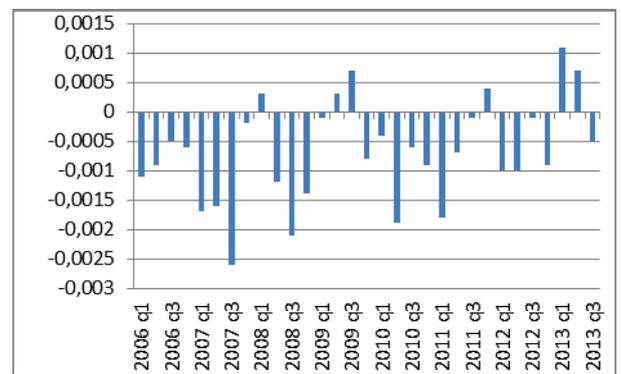
Log city boundry dist.



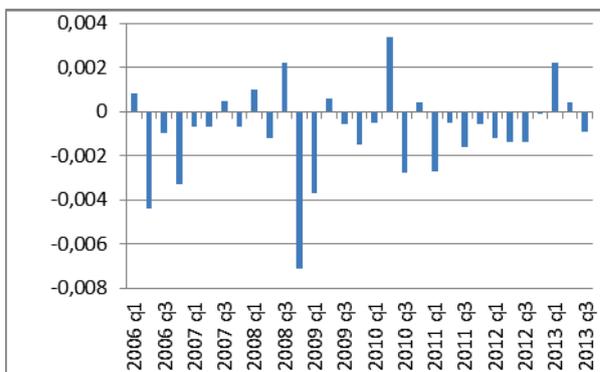
Log metro station dist.



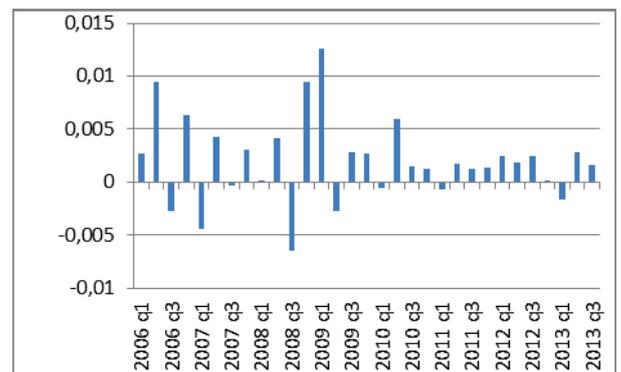
Tram noise



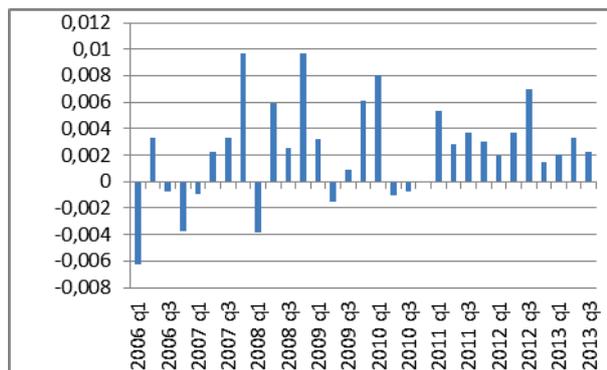
Industrial noise



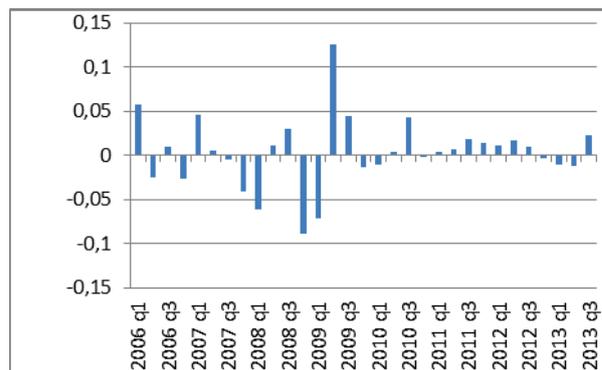
Green areas in 1000 m. circle



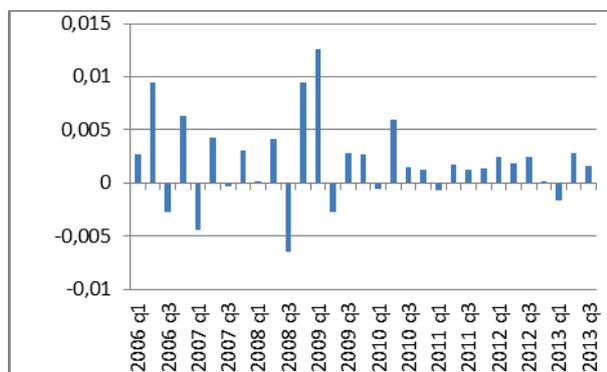
Medical center in 1000 m. circle



Log trade center dist.



Log green area dist.





---

## **19. Monitoring 15 years of residential house price development in Hungary with the help of the FHB House Price Index**

**Gyula Nagy**



dr Gyula Nagy – director of FHB Mortgage Bank Plc, Budapest, Hungary,  
nagy.gyula@fhb.hu

### **Acknowledgment**

Special thanks to :  the exclusive research partner of  
the FHB Index

and to Áron Horváth, PHD-,head of ELTINGA

The views expressed in the article are those of the authors and do not involve the  
responsibility of the Bank.

---

## Article 19 Contents

1. Introduction .....	235
2. Source of transaction data and methodology of the FHB House Price Index .....	236
3. The development of house prices in Hungary in the past 15 years .....	239
Phase 1. "Good old days" 1998–2001 .....	240
Phase 2. The period of fiscal expansion 2001–2003 .....	241
Phase 3. The credit boom period 2004–2007 .....	243
Phase 4. The effects of the financial crisis: The "setback" from 2008 .....	245
4. Mortgage markets in the EU and in CEE between 2008 and 2013 .....	248
5. 2014 – end of the setback period in Hungary? .....	250

## **Abstract**

The working paper presents the development of the housing market in Hungary between 1998 and 2013 through the history of the FHB House Price Index. For computing the FHB House Price Index FHB applied the hedonic method. At its first publication in 2009 the Index was based on actual transaction data of residential real estate collected from the year of 1998 from appr. 1,000,000 residential properties located in appr. 3,200 municipalities. The source of data include the valuation records of FHB Mortgage Bank, as well as the buying - selling transaction database purchased from NAV, the national tax authority. Since its first publication the Index is updated on a quarterly basis. The average index value in 2000 was 100 later it peaked at 200,7 in the first quarter of 2008. The 15 years of the housing price history were divided into 4 significantly different eras. When analysing the development of the housing market, relations between selected macroeconomic and financial environment indicators, money market and credit market indices and other data of the housing market were also taken into consideration.

Keywords: House price index, mortgage credit, housing market

## 1.Introduction

Since the outbreak of the latest financial crisis , the exact measurement of house prices - that are the main collaterals behind mortgage loans - has become of primary importance. It is not an easy task to get exact figures of the residential housing transactions that occurred recently on the market and it is even more difficult to prepare the accurate time series for a longer period from these data.

The most popular housing indexes are already existing for several decades, they can be used as a benchmark for the residential housing market of the relevant country, as e.g. the Nationwide Home Price Index<sup>1</sup> ,that is published in the UK on a monthly basis and the commencement of the data collection goes back to as far as the year 1952 or the Case Shiller Index<sup>2</sup> of S & P in the US ,that is the most widely quoted housing market indicator in the world.

The FHB House Price Index measures the development of Hungarian residential real estate prices. It shows how the value of Hungarian residential properties has changed since 1998. FHB House Price Index is already covering 15 years of house price development and has become an established benchmark for monitoring the house price changes on the Hungarian market.

The index is published quarterly, and starts from the first quarter of 1998. The value of the index is normalised with the average of year 2000, i.e. the average index value in year 2000 is 100. During the development of the Index, we regarded such long-standing, internationally well-known indices as in the EU the house price index of Halifax<sup>3</sup> Nationwide or Hypoport Index<sup>4</sup>.

Being a mortgage bank, FHB tries to analyse together with development of the housing market, the main drivers of the mortgage business, the most important economic indices that may positively or negatively influence the housing trends and

---

<sup>1</sup> <http://www.nationwide.co.uk/hpi/default.htm>

<sup>2</sup> <http://www.spindices.com/index-family/real-estate/sp-case-shiller>

<sup>3</sup> [http://www.lloydsbankinggroup.com/media1/research/halifax\\_hpi.asp](http://www.lloydsbankinggroup.com/media1/research/halifax_hpi.asp)

<sup>4</sup> [http://www.hypoport.com/hpx\\_mean\\_en.html](http://www.hypoport.com/hpx_mean_en.html)

the acting and motivations of the market players (developers, financing institutions, customers etc)

## **2. Source of transaction data and methodology**

The index is based on purchase-and-sell transaction data of residential real estate, and was prepared by processing purchase-and-sell data of more than 1 million residential properties from around 3200 settlements from 1998 until 2013 in Hungary . All observations were gathered during the property valuations made for FHB Bank and a further database purchased from NAV<sup>5</sup> (from Stamp Duty Offices before 2008) served as sources of data. When providing credit for the purchase of real estate, FHB Bank devotes special attention to the registration of features of the properties. During valuation for the assessment of mortgage lending value (MLV, collateral value), several attributes of a property are investigated:

- Address, location
- Area (gross, net, useful)
- Lot size
- Year of construction
- Distance from and access to public transportation (train, local- and distance bus)
- Condition of the property
- Technical features of the construction
- heating system etc.

Another – although less detailed - part of the database was collected by the Stamp Duty Offices. Every single Hungarian residential property transaction must find its way to this database, which comprise the following:

- Date of purchase
- Address (settlement, district if in Budapest)
- Purchase price documented through transaction
- Sum of valuation by the Stamp Duty Office

---

<sup>5</sup> NAV – National Tax Administration Office (APEH before 2010)

- Type of building (detached house, terraced house, condominium and flat in a block of flats etc.)
- Area

The Act XXX. on Mortgage Banks and Mortgage Bond introduced the so called “special bank principle” in Hungary and FHB Bank was the first mortgage bank in Hungary to enter into this business field in 1998. Being a mortgage bank engaged mostly in residential mortgage lending FHB put special emphasis on collecting data of house price valuations for the exact measurement of the mortgage value of properties, that serves as collaterals for the mortgage loans. The law also made it possible to provide property valuation and related services to third parties. (Section 3 § 5 says *Apart from the financial, investment and non-core investment services mentioned in subsection (2), mortgage banks may only engage in the business of providing appraisal services to determine the mortgage lending value of real properties.*) The prevailing legal regulations<sup>6</sup> prescribed very strict valuation criteria for the mortgage banks. Among other methods they prescribe the use of comparative valuation methods, and the application of the principle of carefulness in the valuation process. To comply with these criteria FHB started the collection of the transaction data followings its establishment.

---

<sup>6</sup> -Decree of the Minister of Finance No. 25/1997. on the Calculation Methods of the Mortgage Lending Value of Real Estate

-Decree of the Minister of Agriculture No. 54/1997 on the Calculation Methods of the Mortgage Lending Value of Real Estate Qualifying as Agricultural Land.

## Methodology

For computing the index, FHB applied the process based on the hedonic method, which is usual in the profession, and the classic reference of which is the study of Kain and Quigley<sup>7</sup>. The method serves to filter out the composition effect appearing in observations of simple indicators (mean price, median price). The composition effect itself arises, because not every single property is sold in every quarter, thus our observations – even if the transaction data are complete - are accordingly a restricted sample of the real estate stock, and the sample is not representative. The hedonic method is the most widespread statistical way to measure aggregate property price change, and is based on the theoretical consideration that the value of houses/flats can be divided into the values of their attributes.

The filtering conditions concern that

- the transaction should be between private persons.
- The whole ownership should be transferred.
- An existing date should be connected to the purchase price.
- An area should be belong to the purchase price.
- The price should be reasonable.
- The price of one area unit should be reasonable.

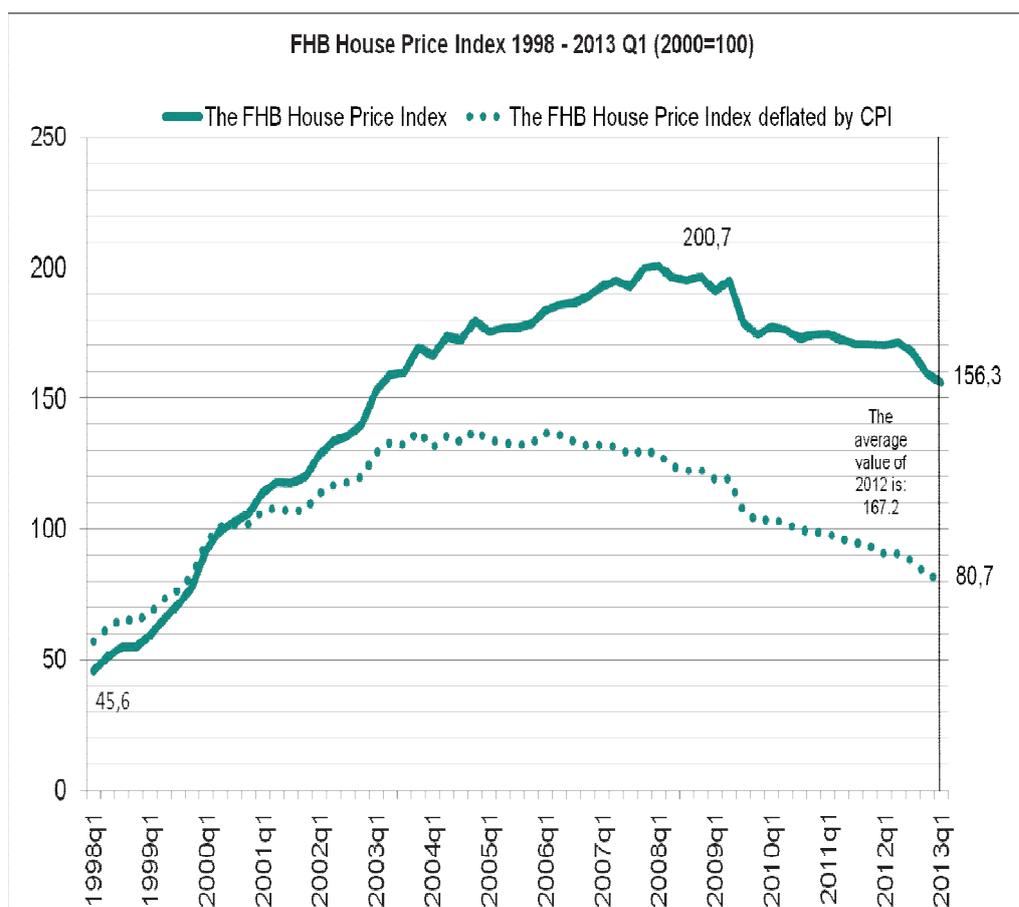
The development of the methodology and the quarterly refreshment of the Index happens through a cooperation the between FHB Mortgage Bank, FHB Real Estate Co. and ELTINGA<sup>8</sup>, a university based real estate research co. To assure the reliable data and to verify the outcome of the research, there are also regular consultations with the Central Statistical Office (KSH) concerning methodology and data filtering issues.

---

<sup>7</sup> Kain, John F. – John M. Quigley [1970]: Measuring the Value of Housing Quality. *Journal of the American Statistical Association* 65 (330), Jun 1970, pp. 532-548.

<sup>8</sup> [www.eltinga.hu](http://www.eltinga.hu)

Fig.1.



Source: FHB

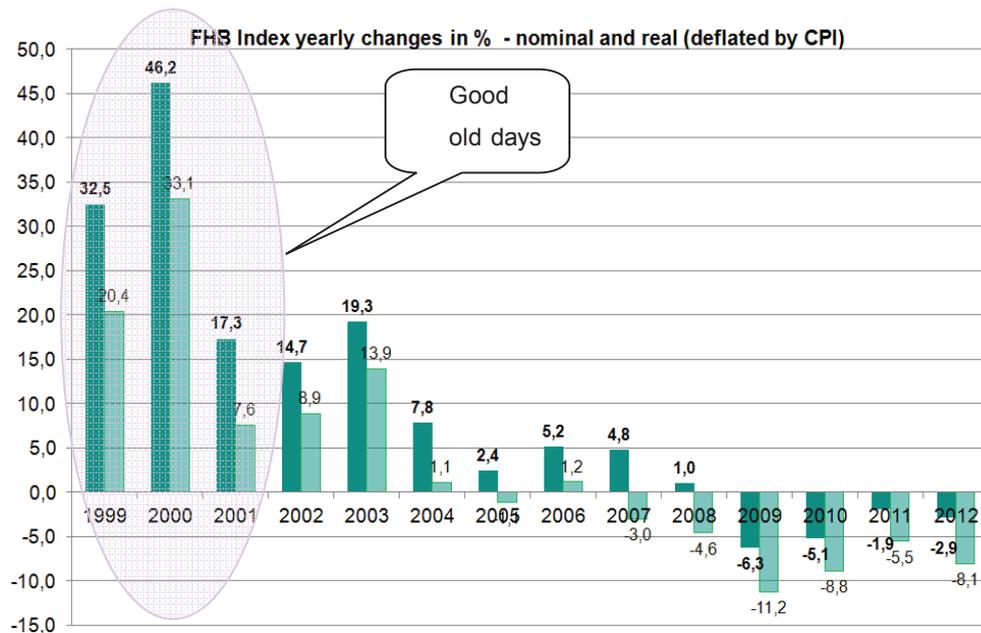
### 3. The development of house prices in Hungary in the past 15 years

At the time of presenting this paper the FHB Index covers already 15 years of the housing price history in Hungary. This period can be divided into 4 significantly different eras. At the beginning of this period prices have been radically increasing both nominally and in real term. During the so called “fiscal expansion” and “credit boom” periods, prices have been increasing in nominal terms, but the stagnation in real term was already apparent. Since 2008, the trend has been clear: nominal housing prices in Hungary were declining, thus, we have been experiencing a downward trend for the last five years. Thanks to this, Hungary now has one of the lowest average prices in the European Union (on average less than EUR 1,000/ square meter) as well as the lowest number of newly built homes. Hungarian house prices are considered low even for the CEE region.

**Phase 1. “Good old days” 1998 -2001**

The first phase, lasting from 1998 to the end of 2000, saw a steep rise in both nominal and real prices. Real estate prices increased by 132% over the three-year period and almost doubled in real terms. Although continually growing demand at the end of the 1990s had not yet met with a corresponding level of supply, the expectations concerning the coming growth of the housing market and the preparation of a new housing subsidy system fostered price increases. The macroeconomic situation also evolved favourably during these years as the economy embarked on a steady growth trajectory. Disposable incomes of private individuals increased by 46% over this period.

Fig.2.



Source: FHB

It is interesting to mention however that the amount of housing credit taken out at the time was negligible, as a total of only HUF 53 billion in home loans migrated to the household sector over the entirety of the period concerned, while interest rates on loans were also relatively high at around 20%. In spite of this, improving income for households and favourable future economic outlook stimulated the demand on the housing market, faced with a supply that was waning in volume and becoming increasingly outdated in quality. In the mid-1990s, in the few years following

budgetary stabilization, the effects of declining demand were also felt on the housing market and the renascent inclination to invest initially appeared mainly on the equities markets. The turning point came with the Russian crisis of 1998 and the stock market crash, in the wake of which many investors sought opportunities on the housing market rather than on the stock exchange. According to data of the National Bank of Hungary (MNB), private individuals invested a total of HUF 108 billion in the stock market in the two years prior to the crash, while some HUF 160 billion was withdrawn from the value of equities in the five quarters following the crash. It is important to note that prior to the crisis, the macroeconomic environment had been developing favourably, with growing incomes and employment levels paving the way for an explosion in demand on the housing market. At the same time, however, supply was outdated and still fell well short of the standard required by demand. The steep rise in housing prices is also partly explained by the still comparatively high, double-digit inflation that characterized Hungary at the end of the 1990s, as consumer prices rose more than 30% over the examined three-year period. At the start of the decade, citizens continued to favour putting their savings into tangible assets in a period of relatively high inflation.

## **Phase 2. The period of fiscal expansion 2001-2003**

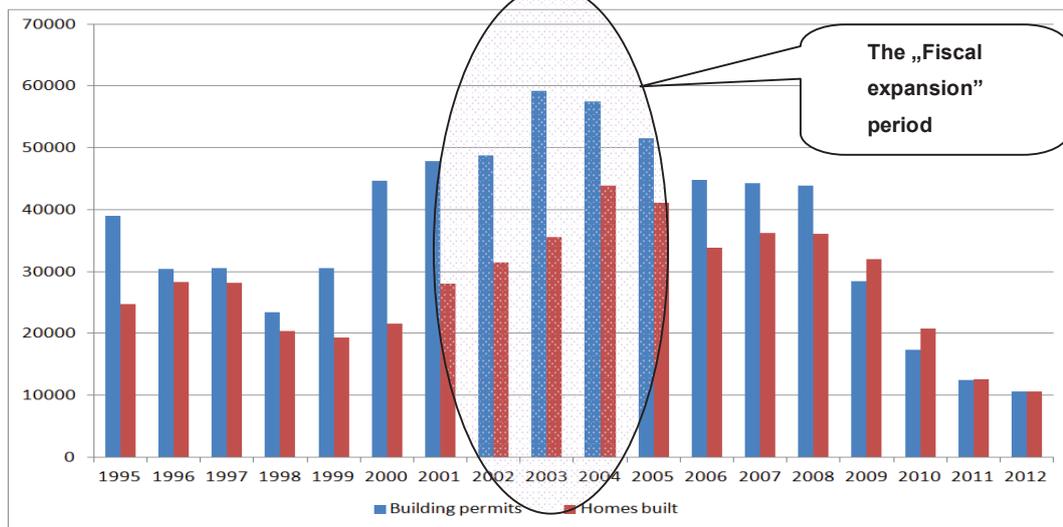
In the period from the beginning of 2001 to the end of 2003, housing prices continued to rise significantly, although more modestly than before, by a total of 60% over the period, representing an annual average increase in real value of 11%. The introduction of the housing subsidy system in 2001 provided fresh impetus to the development of the Hungarian housing market, following the temporary stalling of price increases at the end of 2000. Impacted by the introduction of subsidies, rates of interest on home loans fell substantially lower than market interest rates, thus making this new form of financing attainable to a broad swathe of the population. The launch of subsidies may also have influenced expectations on the supply side in the long run, as it was from this point on that a discernible expansion in supply occurred.

Parallel with the introduction of subsidies, a significant revitalization of the budget occurred from 2002, resulting in a combined 36% increase in household incomes over the three-year period, representing 15% growth in real terms. In addition to rising incomes, a large amount of external resources appeared in the sector on the

credit side, in an amount of approximately HUF 430 billion annually. It is also worth noting that credit growth in these years was driven exclusively by forint loans with subsidized interest rates.

The continued boosting of demand through preferential lending then met with a revival on the supply side from 2001 onwards, with a substantial increase in the number of new homes built . This period saw developers launch the first large-scale residential projects in all major cities in the country. On average over the period, a 50% increase in new homes was added to the housing market annually, amounting to an annual average of 31,703. The number of newly built homes first exceeded the 40 thousand annual units with 43,913 pcs in 2004, while building permits peaked in 2003 with almost 60 thousand permits. Issued building permits then gradually decreased from 2004 and onwards and housing permits(completions) followed the trend with the usual 9 -12 months delay.

Fig.3 Number of building permits and housing permits on yearly basis



Source: KSH<sup>9</sup>

The expansion of supply necessarily occurred only gradually, and for this reason, prices continued to rise until the end of the period. Other amendments to the system of subsidies occurring between 2001 and June 2003 continually eased the conditions for drawing on subsidies, expanding the options for utilizing credit and increasing the size of subsidies available. Among these changes, the most

<sup>9</sup> KSH = Hungarian National Statistical Office

significant step was taken in March 2002, when subsidies were extended to include used homes. The effect - that the introduction and expansion of subsidies had on driving prices up - can be easily traced in the evolution of the FHB House Price Index. However, the generous subsidy system had already placed an unsustainable burden on the budget by 2003, when the first tightening measures were taken, the effect of which began to be felt on the development of housing prices.

### **Phase 3. The credit boom period 2004 -2007**

The tightening of the subsidy system made its effects felt immediately on housing prices as a one time steep rise in prices was checked at the end of 2003. The next few years would be characterized by a considerably more modest annual rate of nominal growth of around 4%, as price increases merely kept pace with – and subsequently fell short of – the rate of inflation.

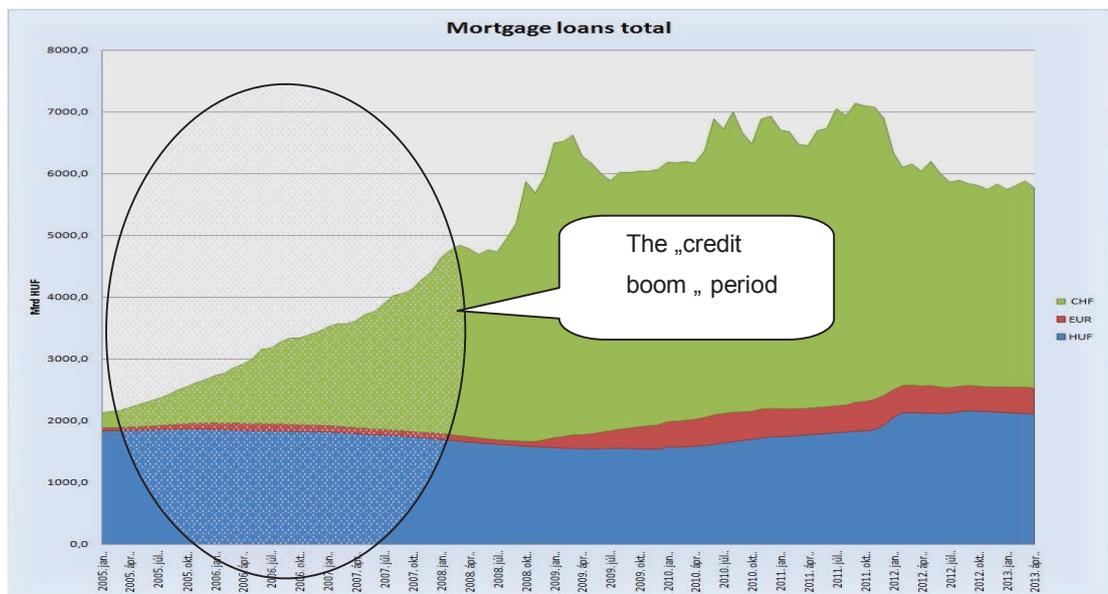
In many respects, this was to prove to be a period of contradictions, replete with inconsistencies and impacts that cancelled one another out. It is true that housing prices across the nation increased by a further 18% over this period, and that it was really due to the accumulation of the budget deficit that the value of homes decreased annually by 1% in real terms on average over this four-year period, alongside spikes in inflation on two occasions (at the beginning of 2004 and the end of 2006). The decline in real values began in smaller fluctuations from the beginning of 2005; however, the majority of households, suffering the “money illusion” of being unable to distinguish nominal from real values, scarcely perceived this depreciation at the time, and even if they did, then presumably they did not believe it to be a long-term effect. Despite the tightening of conditions for subsidies, competition for lending was still able to sustain demand for a number of years, with a similar amount of loans disbursed annually compared to the preceding period. Once subsidies were terminated, lower-interest forex loans (mainly CHF) became popular and households began to build up debts in foreign currency. From 2004, an increasing proportion of mortgage loans were originally denominated in foreign currencies. At the same time, towards the end of this period a growing proportion - more than 50% of mortgage loans in 2007 - were freely usable mortgage equity loans. This clearly demonstrates that private individuals were using a decreasing portion of their borrowed financial resources for the purchase of homes, and an increasing portion for spending on consumer goods or redeeming more expensive

loans taken out earlier, which in turn also tended to indirectly stimulate consumption. The tightening of the subsidy system naturally had its effect on the supply side as well, as the number of issued building permits began to decline from 2004 (Fig. 3).

Certain delayed effects, however, meant that the supply of new homes still continued to expand, by an annual average of 38,755 units over this period. While incomes increased by barely 1% annually over the period, the abundant supply of credit sustained demand, and consequently housing prices did not begin to fall until the beginning of 2008. Classic macroeconomic factors seemingly no longer supported a further rise in housing prices at this time, but the copious supply of credit at the global level concealed these effects. Partly as a result of this, and as a consequence of the concomitant global boom on the housing, massive foreign demand appeared in Hungary, and particularly in Budapest.

In addition to the mainly German and Austrian investors who had already arrived in earlier years, this period saw the appearance of British, Irish, Spanish, and finally, Russian investors. The foreign demand that emerged on the Hungarian market for the most part took the form of purchases for investment purposes, and a proportion of the profits previously realized on other markets now made their way to Hungary

Fig 4. The growth of the total mortgage loan portfolio 2005-2013

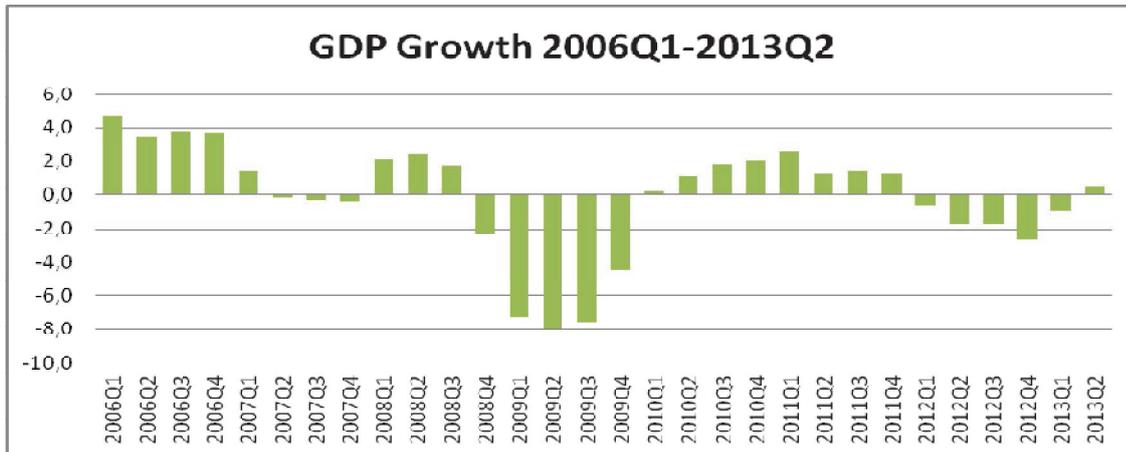


Source: National Bank of Hungary

**Phase 4. The effects of the financial crisis: The “setback” from 2008 .**

The crisis has seriously affected the Hungarian economy, GDP contraction was significant following the outburst of the financial crisis.

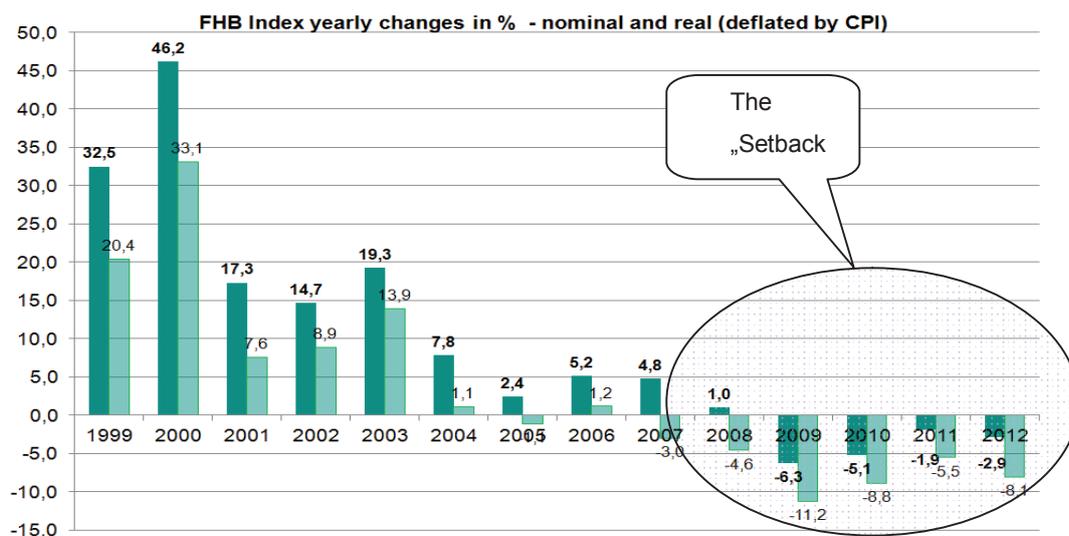
Fig 5.



Source KSH

A long period of rising prices looked set to be broken at the end of 2007, when the factors regarded as the most important from the point of view of the housing market suddenly began to align in the same direction: supply became saturated, demand fell, disposable incomes levelled out and then declined, and more and more news arrived indicating that the great real estate boom abroad had come to a standstill.

Fig.6.

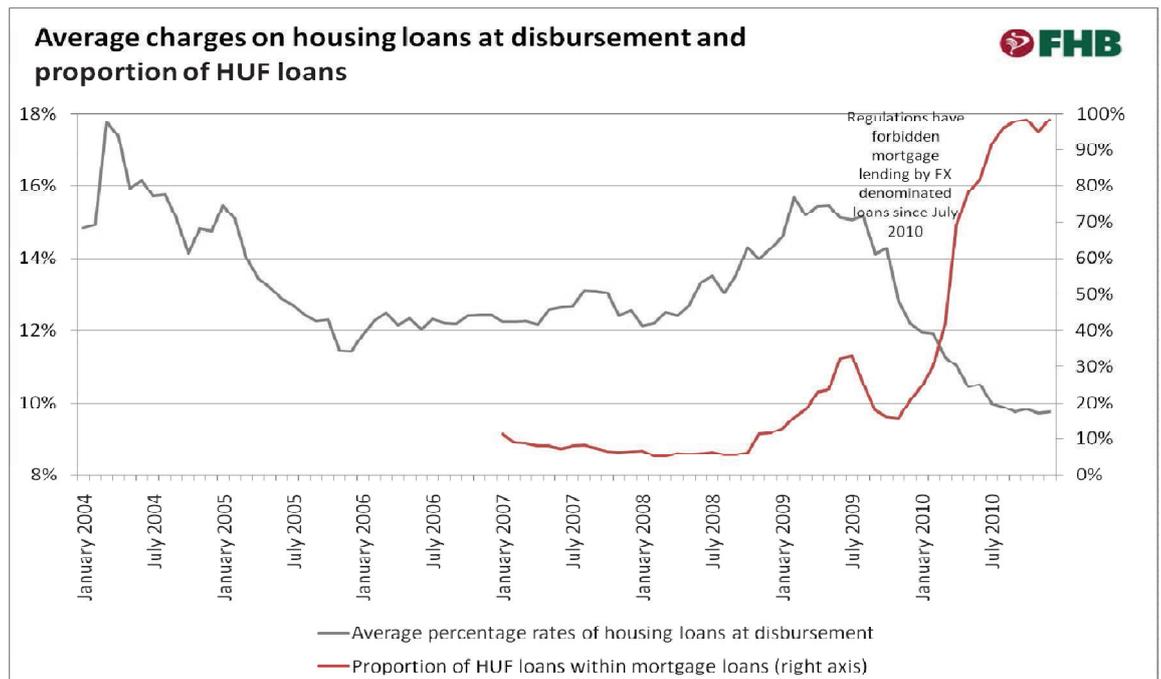


Source FHB

The economic crisis in the so called “setback” period radically affected the demand mainly through two different channels: layoffs and salary corrections resulted in a significant reduction in the real income of the population, while corrective actions taken in the financial industry increased the price of offered loans. The construction of new housing began to fall in 2009, concurrently with the increase in the number of unsold new housing units experienced by the developers.

Changes in regulations caused a turn toward Forint loans (Fig.7). Average percentage rates of new HUF loans were more favourable because lenders were offering products with lower initial interest rates. In 2010 the granting of FX denominated loans was forbidden. As a result, new FX mortgage loans disappeared from the market, although the existing FX stock had an increasing burden for the debtors (appreciation of the CHF versus HUF)

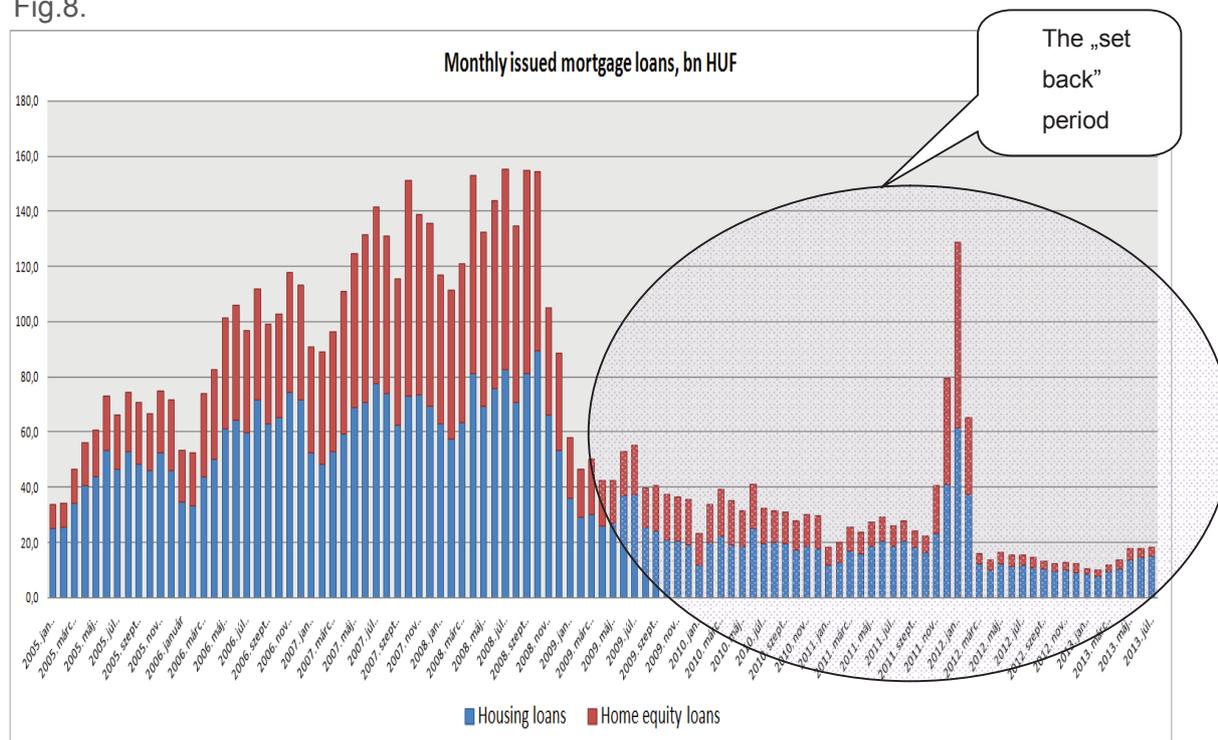
Fig.7.



Source: Hungarian National Bank, FHB

The amount of the newly issued mortgage loans has fallen to 1/5 th of the amount issued before the crisis. (Fig.8.) Following the introduction of the „early repayment scheme”<sup>10</sup> and temporary increase of „remortgage” activity the amount of newly issued mortgage loans has decreased even further.

Fig.8.



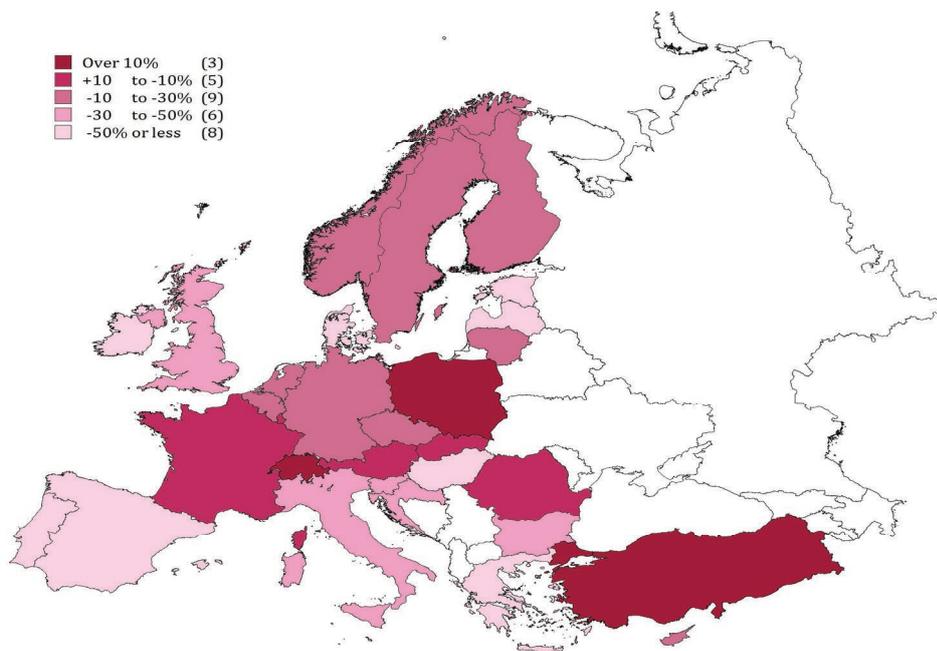
Source:MNB

<sup>10</sup> „In the frame of the „early repayment scheme” mortgage debtors were entitled for a preferential full repayment of EUR, CHF and JPY mortgages according to Section 200/B. of Act CXII of 1996. on Credit Institutions and Financial Enterprises. Following fixed exchange rates were defined 180HUF/CHF, 250 HUF/EUR, 200 HUF/100 JPY.

#### 4. Mortgage markets in the EU and in CEE since 2008

It can be seen that for the whole period between 2008 and 2012 already in several countries an increase can be observed, so the price declines caused by the crisis are no longer felt. In case of Hungary, however, the decrease is continuous.

**Fig 9. The impact of the crisis on new-build homes: average of 2008-2012 period compared to the average of 2003-2007 period**

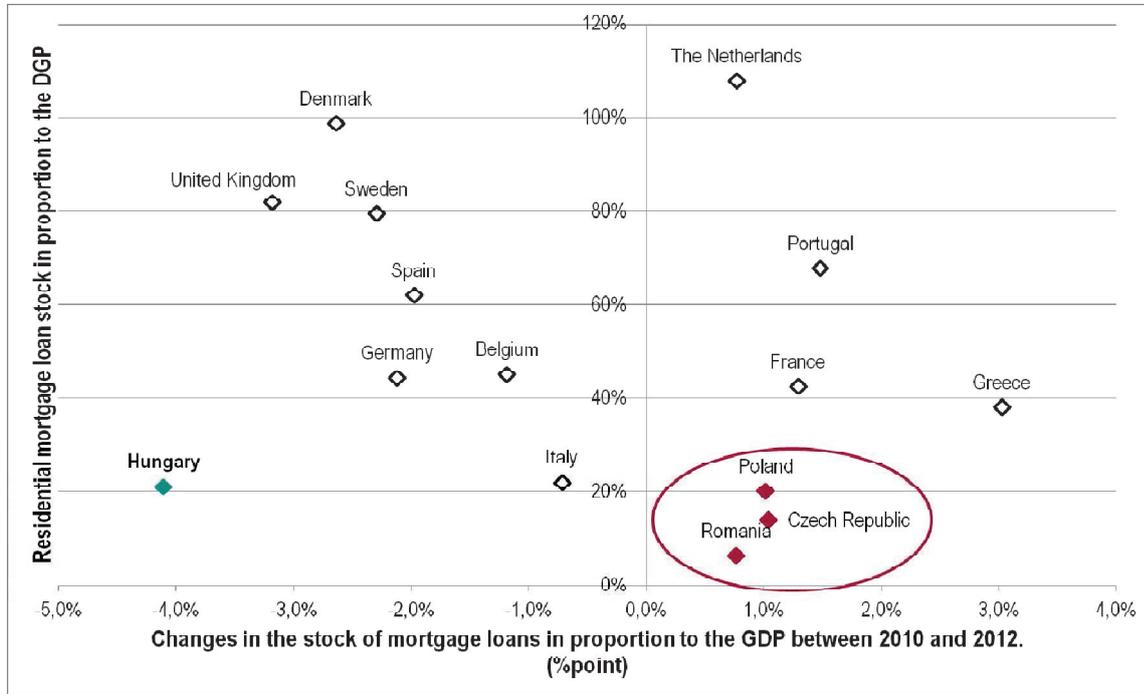


Source: Eurostat

The crisis still affects the housing markets. Fig 9. shows that the exceptions are those countries where the number of new-build housing units increased in the past five years compared to the period before the crisis. Turkey is developing fast, while Poland survived the crisis in a good condition. Several countries' mortgage markets are still characterized by the deleveraging process of households. Fig 10. shows that in the more indebted countries the mortgage loan stock compared to the GDP has rather decreased since 2010. In Hungary as a consequence of the early repayment scheme and the low take out of new mortgage loans the stock dropped in a large extent, while in other countries in the Central and Eastern European region mortgage loan portfolio even increased compared to the GDP.

**Mortgage loans in proportion to the GDP in 2012 and changes happened between 2010 and 2012**

Fig10.

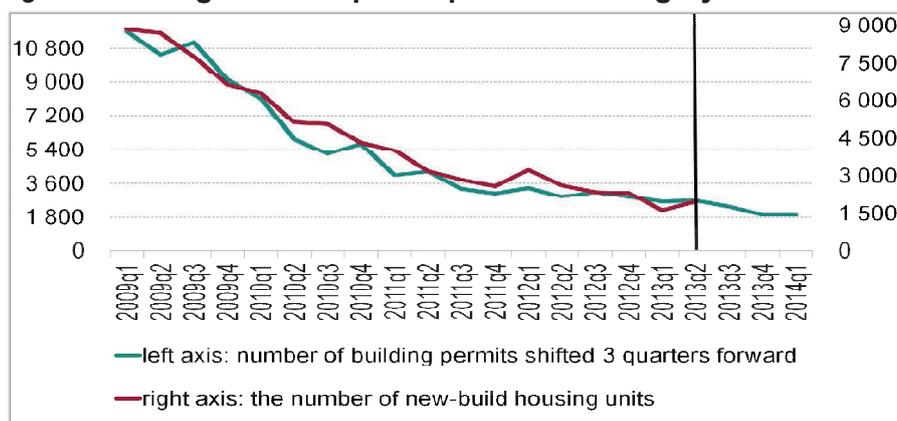


Source: Eurostat

## 5. 2014 – End of the setback period in Hungary?

As far as the supply side is concerned in 2013 the number of new-build houses continued to decrease. Until June 2013 altogether 2680 new-build units were handed over, which is a further decline even if we compare it to the previous year. The number of building permits anticipates the number of future housing completions quite well. , Fig.11 shows the development of housing completions and building permits from 2009 (building permits are shifted 3 quarters forward). Based on this figure we can say that building permits do not give any reason to optimism , their number is still decreasing and in the first half of the year only 3401 permits were issued.

Fig 11. Building and occupation permits in Hungary

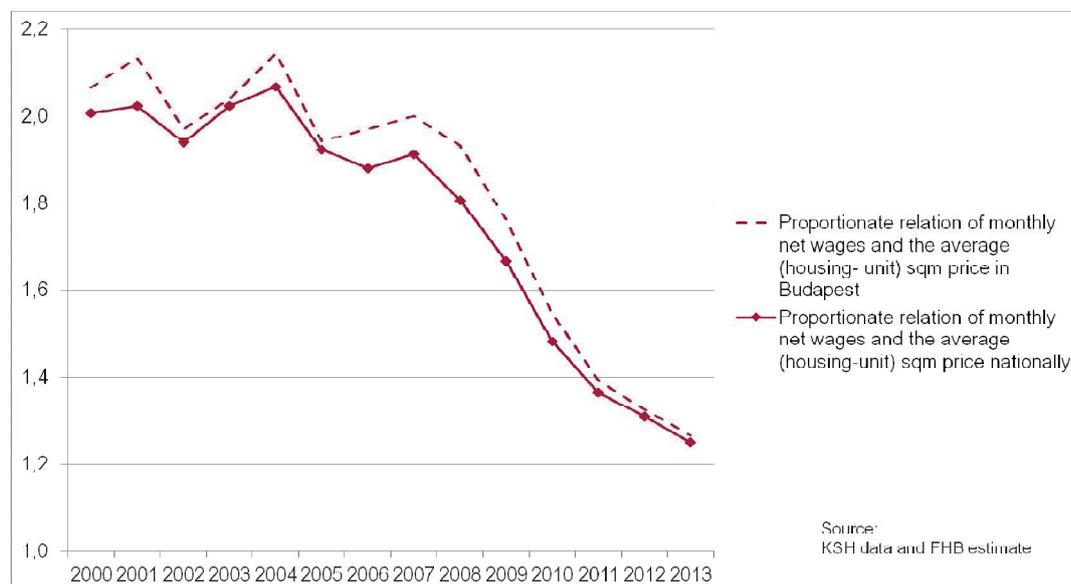


Source: KSH

However there are also positive signs on the market. Currently available housing subsidies are much more favorable than the possibilities of the recent years. According to our calculations, in which we summarized the value of the potential subsidies available in case of a housing loan for a typical house, we may state that one has to go back to year 2004 to find better conditions than the presently available ones. The net monthly earnings of employees increased by 2.6% in the first quarter of the year, which altogether may increase the demand for houses: In 2013 one had to work less for a square metre of a housing-unit than at any time in the past 13 years. (Fig.12)

## Development of housing-unit prices in relation to net wages (Sources: Hungarian Central Statistical Office data and FHB estimate)

Fig 12



As far as the Central Bank base rate is concerned, with 3% at the end of 2013 it reached its lowest value in the last 25 years. According to the National Bank of Hungary forecast, economic growth will also pick up in 2014 in Hungary. (GDP growth of 2% is forecasted by the National Bank of Hungary for 2014)

In parallel with the declining inflation rate, interest rates on deposits and loans are continuously decreasing as well. Subsidized credit products are available with 6-7% initial interest rate (first 5 years period).. After several years of downward trend mortgage lending has moderately increased in the second part of 2013 and is expected to continue to rise at the beginning of 2014.

The above economic indices may positively stimulate the demand side on the mortgage market. This may lead to a turning point point in 2014, although it is too early to draw final conclusions, as the positive signs are still weak and Hungary needs a longer period of positive changes to have lasting effect on the housing and mortgage market.

## References

- Act no. XXX of 1997 on Mortgage Banks and Mortgage Bonds
- 
- Decree of the Minister of Finance No. 25/1997. on the Calculation Methods of the Mortgage Lending Value of Real Estate
- 
- Decree of the Minister of Agriculture No. 54/1997 on the Calculation Methods of the Mortgage Lending Value of Real Estate Qualifying as Agricultural Land.
- 
- Kain, John F. – John M. Quigley [1970]: Measuring the Value of Housing Quality. Journal of the American Statistical Association 65 (330), Jun 1970
- 
- Record low prices on the Hungarian housing market, Resource Magazin 2013. 8-9. p.28-29
- 
- <http://www.fhbindex.com/FHB-Index/downloads/FHB-House-Price-Index/Archive/2009%20First%20publication.pdf>
- 
- <http://www.fhbindex.com/FHB-Index/downloads/Methodology/Methodology%20long%20version.pdf>
- 
- [http://www.fhbindex.com/FHB-Index/downloads/FHB-House-Price-Index/Latest-issue/FHB\\_Index\\_2013q1\\_eng%20\\_2\\_.pdf](http://www.fhbindex.com/FHB-Index/downloads/FHB-House-Price-Index/Latest-issue/FHB_Index_2013q1_eng%20_2_.pdf)
- 
- Hungarian National Bank statistical data [http://english.mnb.hu/Statisztika/data-and-information/mnben\\_statisztikai\\_idosorok](http://english.mnb.hu/Statisztika/data-and-information/mnben_statisztikai_idosorok)
- 
- KSH - National Statistical Office [http://www.ksh.hu/housing\\_and\\_public\\_utilities](http://www.ksh.hu/housing_and_public_utilities)
- 
- For mortgage loans in GDP proportion for the differen EU countries <http://www.hypo.org/DocShareNoFrame/docs/1/HHKLGAGDFJEIENAFNOONNPBBPDW69DBDA6TE4Q/EMF/Docs/DLS/2013-00033.pdf>
- 
- MNB, The baseline inflation and GDP projection of the Inflation Report, September 2013

---

## 20. Logit modelling as a tool supporting decision making in the real estate market

**Łukasz Mach**



dr eng. Łukasz Mach – Narodowy Bank Polski, Regional Branch in Opole, Department  
of Statistics and Foreign Exchange

---

## Article 20 Contents

1. Introduction .....	257
2. The research process- grouping provinces into homogenous area .....	258
3. The research process – application of logistic regression in specified groups of provinces .....	261
4. Summary .....	269
Literature .....	270

**Abstract:** The article includes a logit model construction process to support the process of decision making in real estate market. The model of logistic regression which has been elaborated, shall define the probability of transactions in the real estate market and it will indicate statistic variables which influence the demand significantly. The process of decision making (based on logit models) - substantially prepared and correctly executed - is a key determinant having influence on improving the competitiveness of companies, especially during the global economic crisis.

**Keyword:** logit modelling, residential real estate market, decision making process

## 1. Introduction

Essentially prepared and properly conducted decision making process is the key factor, which influences the competitiveness of companies. Well prepared decision making process enables to minimize the risk of making wrong decisions and helps to enhance the competitive advantage of companies. During the economic crisis, especially companies, which operate in the prone to crisis industries should consider the proper decision making process. Taking into consideration economic theory, the main factors, which influences on competitiveness of economies there are investments implied mainly by housing market. The growth dynamics can be parameterized by the volume of sold residential real estates.

The aim of the article is researching the secondary real estate market in Poland by building three qualitative decision making models. The aim of the above mentioned is estimation of probability of selling real estates and indication (for every single model) diagnostic variables, which characterize particular residential real estates and influence on selling them. Building three different models was a result of analysis of residential real estate on the homogenous markets i.e. the markets homogenous according to the assumptions.

The process of grouping residential real estate markets in provinces for homogenous groups was conducted by applying cluster analysis (by using inter alia macroeconomic data of economic and business data)<sup>1</sup>. The research process, which aims at building quality models describing probability of selling residential real estate based on the model of logistic regression (with the use of bidding data and selling data from third quarter of 2013 in capital of provinces<sup>2</sup>).

---

<sup>1</sup> data from Central Statistical Office

<sup>2</sup> data from National Bank of Poland

## **2. The research process- grouping provinces into homogenous area**

The research process was conducted in two main stages. In the first stage, taking into consideration socio-economic factors which affect substantially on the price of square meter of residential, Polish provinces were grouped into three homogenous groups. Determinants describing residential real estate and their influence on selling apartments were specified in the second stage by applying logistic regression. In the process of grouping, by applying cluster analysis technique, the key factor was a choice of discriminating variables, which were used for creation of homogenous provinces groups. The assumption, that socio-economic variables would be determinants used in grouping particular districts, was made. The recognition of the factors, which mainly influence on the Polish districts development in the socio-economic dimension, was made with the use of the factor analysis. The observation matrix (16 provinces, 18 variables) was the basis for conducting factor analysis.

Diagnostics variables included: X1- new entities of national economy for 1000 population; X2- entities of national economy, which finished their activity for 1000 population; X3- number of micro enterprises for 10000 population; X4- number of small enterprises for 10000 population; X5- number of medium-sized enterprises for 10000 population; X6- number of large enterprises for 10000 population; X7- Gross domestic product per 1 person; X8- average monthly gross salary; X9- the average pension outside agricultural social security system; X10- retail per 1 inhabitant; X11- average monthly expenses per one person for use of apartment; X12 – average monthly expenses per one person per equipment of apartment; X13- average monthly income per person in household; X14- number of rooms for 1000 population; X15- the average usable area per one inhabitant; X 16- apartments for 1000 population; X 17- apartments for 1000 marriages; X18- new apartments in new residential areas (with the right of usage in the whole building or in particular parts) or in non-residential areas for 1000 population.

Implementation and assumptions referring to the correctness of the calculation in the research process were described in the previous thesis of the author [compare Mach 2012, pages 106-116]. In the above mentioned thesis diagnostic variables, which substantially affect for the price of the square meter of residential, were specified. In the above mentioned thesis, the socio-economic variables, which affect on the price of square meter of residential were specified by applying factor analysis and multiple regression. The factor analysis proved to be an effective tool enabling for detection of hidden development factors. Use of the multiple regression enabled for identification of factors substantially influencing on the price of square meter of residential. Parameterization of the researched relations, which specify development

of particular provinces and indication of their affection for formation of the square meter's price of residential, can help enterprises, which operate in real estate industry as a prerequisite decreasing the risk of decisions, which were made in the managing process.

Finally, the variables, which substantially affected on the price of square meter of residential were economic and business variables including: number of micro enterprises for 10000 population, number of small enterprises for 10000 population, number of medium-sized enterprises for 10000 population, number of large enterprises for 10000 population, Gross domestic product per 1 person, average monthly gross salary, retail per 1 inhabitant, average monthly expenses per one person for use of the apartment, average monthly income per person in household. Diagnostic variables defined as above, were subsequently used as input variables applied in cluster analysis aimed at grouping Polish provinces for homogenous provinces. Ward's Algorithm was used as a agglomeration method and the distance measure was Euclidean distance, while applying clustering analysis. Requirements referring to application of clustering analysis and formal record can be found in thesis of the following authors [Aczel, A.D., 2000, s. 849-916; Panek, T. 2009, s.105-169; Witkowska. D., 2002, s.80-90]. Results of applying cluster analysis are depicted at the illustration 1.

Binding distance between 16500 and 48000 was made as an assumption during the division of the Polish provinces into three homogenous groups. For the reason of the article groups were defined as follows:

- Group 1- provinces with small local real estate markets;
- Group 2- provinces with medium and big local real estate markets;
- Group 3- provinces playing vital role in the growth of the residential real estate market.

The following provinces were classified to the group 1: lubelskie, podkarpackie, podlaskie, świętokrzyskie, warmińsko-mazurskie, lubuskie, opolskie, zachodnio-pomorskie, kujawsko-pomorskie. To the group 2: łódzkie, pomorskie, dolnośląskie, małopolskie, wielkopolskie, śląskie. To the group 3 only mazowieckie province.

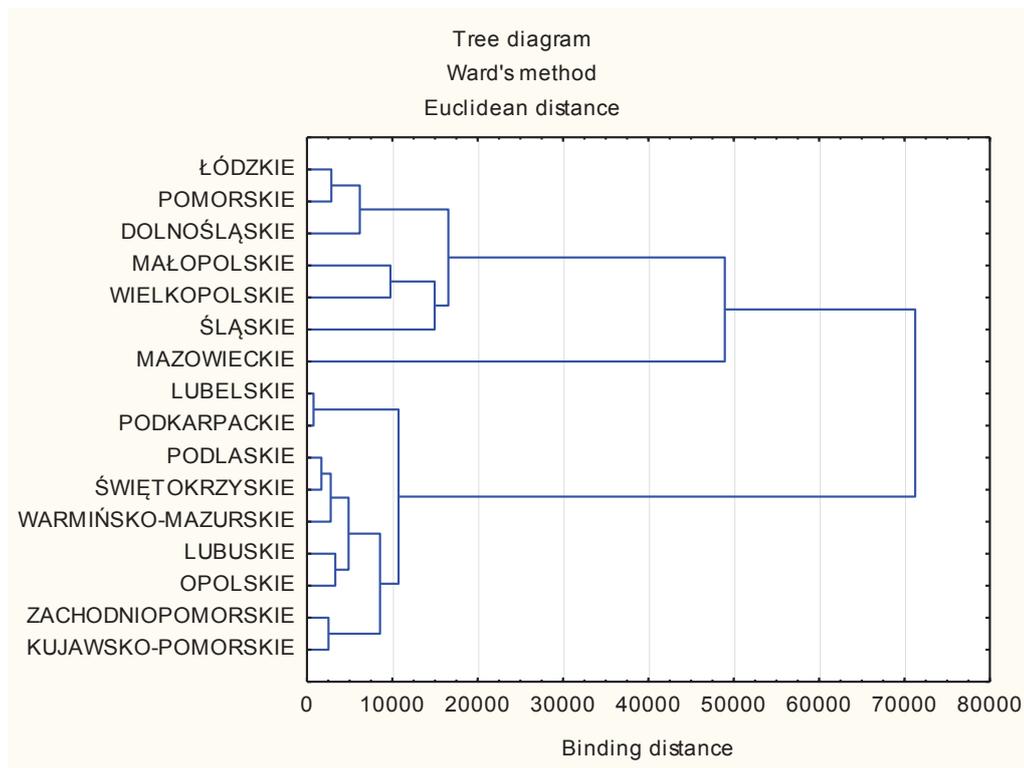


Illustration 1. The result of applying cluster analysis in order to distinguish similar provinces

### 3. The research process – application of logistic regression in specified groups of provinces

The first activity during the creation of logit models was checking whether outliers exist. Subsequently the aim of the research was defined: building a logistic model, which parameterizes the probability of selling the apartment according to qualities by which an apartment is characterized. Input logistic regression model was created for all groups of provinces and then defined in the formula 1:

$$P(Y = 1 / x_1, x_2, \dots, x_9) = \frac{e^{a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + a_5 x_5 + a_6 x_6 + a_7 x_7 + a_8 x_8 + a_9 x_9}}{1 + e^{a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + a_5 x_5 + a_6 x_6 + a_7 x_7 + a_8 x_8 + a_9 x_9}} \quad (1)$$

where:

$x_1$  - the number of floors in the building (building with 5 floors at maximum, building with more than 5 floors);

$x_2$  - number of rooms in the apartment (one-room apartments, two-room apartments or three-room apartments);

$x_3$  - kind of the kitchen in the apartment (dark, bright, annex);

$x_4$  - estimation of location (bad location, average, good);

$x_5$  - estimation of apartment's location in the building (not very favourable, average, favourable);

$x_6$  - floor, at which room is located (first floor, ground floor or the top floor, middle floor);

$x_7$  - the area of the apartment (apartment up to 40 m<sup>2</sup>; apartments with area from 40 m<sup>2</sup> to 80 m<sup>2</sup>, apartments with the area of more than 80 m<sup>2</sup>);

$x_8$  - standard of interior completion (high standard, average, low);

$x_9$  - building technology (traditional, traditional improved, monolithic, prefabricated, wooden, steel frame);

$a_0, a_1, a_2, \dots, a_9$  – structural parameters of the model.

After defining formal logit model (compare formula 1) the estimation procedure was conducted three times in sequence, then verification of parameters of structural models for provinces from group 1, group 2 and group 3 and results interpretation was made. Requirements referring to application of logistic regression and formal record can be found in

thesis of the following authors [Dittmann, P., 2004, s.137-138; Maddala, G.S., 2008, s. 371-382].

### 3.1. Use of logistic regression for provinces from group 1- estimation, model verification and interpretation of result

Nine provinces were classified to the first group. The smallest one is opolskie province and the biggest one lubelskie province.<sup>3</sup> Shortly characterizing secondary residential real estate market we can state as follows:

- There were 4881 records in the first group, where condition of the real estate market was described with 449 selling transactions. In order to precisely classify cases in the created logistic model, the file with 899 was created (with sustainable structure of offers and transactions, where accordingly 450 offers and 449 transactions were located);
- In the first group of provinces, the lowest price for the square meter was 2174 PLN and the highest 6527 PLN;
- 72 % of apartments were located at the buildings with maximum 5 floors;
- 74% of apartments had 2 or 3 rooms;
- 80% of apartments had bright kitchen;
- 39% of apartments were located on the ground floor or on the top floor.

In table 1 preliminary estimation results of logit model were presented with the use of Quasi-Newton method. The aim of the above mentioned was estimation of probability of selling residential real estate located in provinces from group one.

Table 1. Preliminary results of estimation logit model (own study)

	Const	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>
Estimation	-2,07	-0,21	-0,35	0,87	-0,31	-0,27	-0,16	0,69	0,19	0,50
Standard deviation	0,13	0,12	0,17	0,37	0,29	0,15	0,13	0,12	0,13	0,11
t(110)	-15,51	-1,75	-2,06	2,37	-1,08	-1,81	-1,23	5,58	1,42	4,38
Standard p	0,00	0,08	0,04	0,02	0,28	0,07	0,22	0,00	0,16	0,00

<sup>3</sup> Data included the actual population of the particular provinces (stated on 31th December of 2012)

Analyzing presented results, variables, which affect substantially for variable  $Y$  are  $x_2, x_3, x_6, x_7, x_9$ . After subsequent rejecting of variables, which are irrelevant statistically, the final version of logit model was expressed in the formula 2:

$$P(Y) = \frac{e^{-0,03-0,57x_2-0,34x_6+0,57x_7+0,70x_9}}{1 + e^{-0,03-0,57x_2-0,34x_6+0,57x_7+0,70x_9}} \quad (2)$$

where:

$x_2$  - the number of rooms in apartment (apartments with more than 3 rooms);

$x_6$  - floor, where the apartment is located (ground floor or the top floor);

$x_7$  - the area of the apartment (apartments to 40 m<sup>2</sup>);

$x_9$  - building technology (prefabricated);

Summary of the estimation process for logistic regression model are presented in table 2. In this table variables list is presented, coefficients for structural parameters, their estimation errors and the marginal error.

Table 2. Results of estimation process for group 1 of provinces significant variables (own study)

	<i>Coefficient</i>	<i>Standard deviation</i>	<i>z</i>	<i>The marginal effect</i>
const	-0,0372132	0,123053	-0,3024	
x2	-0,568865	0,205495	-2,7683	-0,140339
x6	-0,348923	0,144945	-2,4073	-0,0870095
x7	0,56663	0,175294	3,2325	0,139488
X9	0,698333	0,148452	4,7041	0,172085

The statistical value of p for the whole model was lower than 0,05, what confirms the relevance of the model comparing to the model with intercept only. This datum confirms of model creation purposefulness because implies the statement, that created model implements something new. Also interpretation of log-likelihood, which is the measure of fitting the whole model was made. This logarithm is calculated with the use of statistics  $-2 \log$  with maximum likelihood of created model and in model with the intercept only (in the created model accordingly 1024,57 and 1098,18). At the base of the above mentioned values pseudo  $R^2$  was calculated and it equals 0,06.

Interpreting the results (compare table 2) we can draw the following conclusions:

- probability of selling apartment with more than 3 rooms is about 0,14 lower than 1, 2 or 3-room apartments;
- probability of selling an apartment located at the ground floor or the top floor is lower about 0,09 comparing to apartments located at the other floors;
- probability of selling an apartment with less than 40 square meters is 0,14 times higher than selling apartment with more than 40 square meters;
- probability of selling apartment located in the building, which was built in prefabricated technology is about 0,17 higher than buildings built in other technologies.

In table 3 both correct and incorrect classified cases for the model were presented. The odds ratio, which equals 2,59 was calculated (ratio of quotient of correctly classified cases to quotient incorrectly classified). The value higher than unity indicates, that this classification is better than this, which will be conducted by random.

Table 3. Table relevancy (own study)

	Expected 0	Expected 1	Percent accuracy
0	168	194	46,40884
1	109	326	74,94253

### 3.2 The use of logistic regression for provinces in group 2- estimation, verification of models and interpretation of results

Six provinces were classified to the second group. The smallest one is pomorskie province and the biggest one śląskie province.<sup>4</sup> Shortly characterizing secondary residential real estate market we can state as follows.

- There were 7220 records in the second group, where condition of the real estate market was described with 415 selling transactions. In order to precisely classify cases in the created logistic model, the file with 830 records was created (with sustainable structure of offers and transactions, where accordingly 415 offers and 415 transactions were located);
- In the second group of provinces, the lowest price per square meter was 2500 PLN and the highest 12068 PLN;

<sup>4</sup> Data included the actual population of the particular provinces (stated on 31th December of 2012)

- 72 % of apartments were located in buildings with maximum 5 floors;
- 77,5% of apartments had 2 or 3 rooms;
- 70% of apartments had bright kitchen;
- 41% were located on the ground floor or on the top floor;
- 34% of apartments were built in prefabricated technology.

In table 4 were shown results of the estimation of logit model with applying Quasi- Newton method. Preliminary results of logit model estimation (own study)

Table 4. Preliminary results of logit model estimation (own study)

	Const	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>
Estimation	0,11	-0,23	-0,77	-0,42	-0,08	-0,08	0,10	0,72	1,42	0,48
Standard deviation	0,21	0,20	0,26	0,43	0,62	0,24	0,19	0,21	0,36	0,19
t(110)	0,50	-1,15	-2,92	-0,96	-0,13	-0,35	0,51	3,35	3,94	2,57
Standard p	0,62	0,25	0,00	0,34	0,89	0,73	0,61	0,00	0,00	0,01

Analysing the results we can notice that variables  $x_2, x_7, x_8, x_9$  substantially affect for variable Y. After subsequent rejecting of variables, which are irrelevant statistically, the final version of the logit model was expressed in the formula 5

$$P(Y) = \frac{e^{-0,27-0,68x_2+0,71x_7+1,54x_8+0,68x_9}}{1 + e^{-0,27-0,68x_2+0,71x_7+1,54x_8+0,68x_9}} \quad (5)$$

where:

$x_2$  - number of rooms in the apartment (apartments with more than 3 rooms);

$x_6$  - floor, where the apartment is located (ground floor or the highest floor);

$x_7$  - area of the apartment (floors to 40 m<sup>2</sup>);

$x_8$  - standard of apartment's interior finishing (low);

$x_9$  - building technology (prefabricated).

Summation of the estimation process of logistic regression model was depicted in table 5. In table 5, list of variables, structural parameters coefficients, estimation error and the marginal effect are presented.

Table 5. Estimation process results for relevant variables (own study)

	<i>Coefficient</i>	<i>Standard deviation</i>	<i>z</i>	<i>The marginal effect</i>
const	-0,26751	0,107728	-2,4832	
X2	-0,680145	0,259629	-2,6197	-0,168284
X7	0,710262	0,195656	3,6302	0,169132
X8	1,54336	0,358479	4,3053	0,316343
X9	0,683677	0,166316	4,1107	0,165505

Likewise in the previous model, the statistical value of p for the model was under 0,05 what confirms the relevance of the model comparing to the model with intercept only. Pseudo R<sup>2</sup> was at the level of 0,08. Interpreting results we can withdraw the following conclusions:

- probability of selling an apartment with more than 3 rooms is lower about 0,17 comparing to one-room apartment, two-room apartment or three-room apartment;
- probability of selling an apartment with less than 40 square meters was 0,17 times higher than the apartment with more than 40 square meters;
- probability of selling an apartment with low interior finishing is 0,32 times higher than apartment with medium and high standard;
- probability of selling an apartment located in the building built with the use of prefabricated technology is about 0,17 higher than apartments build in another technologies.

In the table 6 both correct and incorrect classified cases for the model were presented. The odds ratio, which equals 2,66 was calculated (ratio of quotient of correctly classified cases to quotient incorrectly classified). The value higher than unity indicates, that this classification is better than this, which will be conducted by random.

Table 6. Table relevancy (own study)

	Expected 0	Expected 1	Percent accuracy
0	166	123	57,43945
1	132	260	66,32653

### 3.3. The use of logistic regression for provinces classified to group 3- estimation,

After applying cluster analysis only one province i.e. mazowieckie was classified to the third group. Defining the secondary real estate market in the third group, we can state as follows:

- 5922 of records in third group and 333 of transactions. Taking into consideration precise classification of cases during the process of building logistic model the file with 666 records was created, with accordingly 333 offers and transactions;
- In the third group of provinces, the lowest price per square meter was 3401 PLN and the highest 15000 PLN<sup>5</sup>;
- 54% of apartments were located in the buildings with maximum 5 floors;
- 75,6% of apartments had 2 or 3 rooms;
- 60% of apartments had bright kitchen;
- 36,8% of apartments were located on the ground floor or on the top floor;
- 25% of apartments were built with the use of prefabricated technology.

In table 7 were shown results of the estimation of logit model with applying Quasi- Newton methods.

Table 7. Preliminary results of logit model estimation (own study)

	Const	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>
Estimation	-0,08	0,38	-0,12	-0,19	-0,33	0,43	0,27	-0,02
Standard deviation	0,16	0,18	0,28	0,36	0,18	0,22	0,26	0,20
t(110)	-0,53	2,12	-0,44	-0,52	-1,89	1,99	1,07	-0,11
Standard p	0,59	0,03	0,66	0,60	0,06	0,05	0,29	0,91

Analysing the results we can notice that  $x_1, x_7$  substantially affect for variable Y. After subsequent rejecting of variables, which are irrelevant statistically, the final version of logit model was expressed in the formula 6:

$$P(Y) = \frac{e^{0,11-0,40x_1+0,48x_7}}{1 + e^{0,11-0,40x_1+0,48x_7}} \quad (6)$$

where

$x_1$  - number of the floors in the building (building up to 5 floors);

$x_7$  - the area of the apartment (apartment up to 40 m<sup>2</sup>).

<sup>5</sup> Maximum value is the result of the assumption

Summing up the estimation process of the logistic regression model was show in table 8. In this table the list of variables, coefficients for the structural parameters, estimation errors and the marginal effect were presented.

Table 8. Estimation process results for relevant variables (own study)

	<i>Coefficient</i>	<i>Standard deviation</i>	<i>z</i>	<i>Marginal effect</i>
const	0,111986	0,125753	0,8905	
X1	-0,402567	0,157786	-2,5514	-0,100295
X7	0,479893	0,190048	2,5251	0,118881

The statistical value of p for the model was 0,00078 what confirms the relevance of the model comparing to the model with intercept only. What is more it is the proof, that model brings new conclusions. Pseudo  $R^2$  was at the level of 0,02.

Interpreting data we can draw the following conclusions.

- probability of selling an apartment located in the building with maximum 5 floors is about 0,10 lower than in the building with more than 5 floors;
- probability of selling an apartment with less than 40 square meters is 0,12 times higher than selling apartments with more than 40 square meters.

In the table 9 based both correct and incorrect classified cases for the model were presented. The odds ratio was calculated (ratio of quotient of correctly classified cases to quotient incorrectly classified) as 1,44.

Table 9. Table relevancy (own study)

	Expected 0	Expected 1	The percentage of correctness
0	162	171	48,64865
1	132	201	60,36036

#### 4. Summary

The first stage during attempt of estimating the probability of selling residential real estates was the division of Polish regions for relatively homogenous areas. The criterion of division was defined by the set of economic and business variables, which affect significantly for the price of square meter in the residential real estates. During the research three groups of provinces were specified (group 1, group 2 and group 3). Subsequently, basing on the application of logistic regression tools, the attempt of estimation of these three models was made. Models define the probability of selling residential real estate according to the qualities by which are characterized. In the first group of provinces, from 9 diagnostic variables, 4 variables were substantial for estimation the probability of selling: number of rooms (apartments with more than three rooms); the floor, where the apartment is located (ground floor or the top level); area of the apartment (floors till 40 m<sup>2</sup>), standard of the apartment (low), building technology (prefabricated). In the second group, the substantial variables turned out to be number of rooms (apartments with more than three rooms); the floor, where the apartment is located (ground floor or the top floor); area of the apartment (floors up to 40 m<sup>2</sup>), standard of the apartment (low), building technology (prefabricated). Only two variables i.e. the number of floors (buildings with maximum 5 floors), area of the apartment (floors up to 40 square meters) in the logit model were classified to third group of provinces. Each of the three logit functions (compare formula 1,2,3) enables to define the probability of selling residential stocks depending on individual qualities. Moreover calculated marginal effects, give interpretative possibilities referring to residential stocks within provinces included to the particular group. For every estimated logit model table relevancy was appointed and the value of pseudo R<sup>2</sup> was calculated. Taking into consideration results from table relevancy we can state, that correctness of classification of particular models is estimated accordingly at the level of 62,0%; 62,5% and 54,5%. The disturbing fact, that should be noticed, is the very low value of coefficient pseudo R<sup>2</sup>. The process of improvement of input data quality, which characterize secondary residential real estate market, should be continuation of the conducted research. Improvement of data quality would be at the same time improvement of measurers, which describe quality of the models (inter alia pseudo R<sup>2</sup>). Improvement of data quality could be reached through building logistic models considering data from longer period.

## Literature

1. Aczel, A.D., Statystyka w zarządzaniu, Wydawnictwo Naukowe PWN, Warszawa, 2000.
2. Dittmann P., Prognozowanie w przedsiębiorstwie. Metody i ich zastosowanie, Oficyna Ekonomiczna, Kraków, 2004.
3. Mach, Ł., *Determinanty ekonomiczno-gospodarcze oraz ich wpływ na rozwój rynku nieruchomości mieszkaniowych*, *Ekonometria*, 4(38)/2012, ISSN 1507-3866, s. 106-116.
4. Maddala G.S., *Ekonometria*, Wydawnictwo Naukowe PWN, Warszawa 2008.
5. Panek, T., Statystyczne metody wielowymiarowej analizy porównawczej, Oficyna Wydawnicza SGH w Warszawie, Warszawa, 2009.
6. Witkowska D., *Sztuczne sieci neuronowe i metody statystyczne*, Wydawnictwo C.H. Beck, Warszawa, 2002.



---

[www.nbp.pl](http://www.nbp.pl)

