Housing Demand and Demographic Trends: Evidence from Hungary

by

Miklós Farkas

Submitted to

Central European University

Department of Economics

In partial fulfilment of the requirements for the degree of

Master of Arts in Economics

Supervisor: Professor Robert Lieli

Budapest, Hungary

2011

Abstract

This paper investigates the demand for housing in Hungary, focusing on the consequences of a contracting age structure. Demographic trends can influence demand in two ways, namely, by changing total population and by changing age structure. The study first introduces the baseline model of housing demand and then estimates it using panel econometric techniques. Age structure is proxied by the share of young adults in the population and is added to the model as a demand shifter. The results suggest that the price and income elasticities of housing demand are at least one third (in absolute value), while even cautious estimates suggest that shrinking population and aging will, *ceteris paribus*, cause real house prices to decrease by 10% in the present decade.

Acknowledgements

I would like to thank Áron Horváth for introducing me into the topic of real estate economics. Comments from Gábor Kézdi and Balázs Muraközy on my term paper for Econometrics 2 were indispensable for starting off. Last but not least, I express my gratitude to my supervisor, Robert Lieli for his endless patience in helping me to improve my thesis.

Table of Contents

List of Tables

List of Figures

1 Introduction

As a house or a flat constitutes the largest single share of most households' wealth, changes in real house prices should greatly influence the decisions individuals make when choosing how much to consume, how much to work or even choices involving the marriage market. It is also important to keep in mind that different groups of the population often react in opposite ways and by different order of magnitudes to the same house price shock. For example, old homeowners may be forced to substantially decrease their consumption if house prices depreciate, while young renters are better off due to lower housing costs. Therefore, it is vital for economists and policy makers to understand the determination of house prices.

In the short run, demographic trends are not the most important determinants of the residential real estate market, and Hungary is no exception. Since the millennium, Hungary has a well functioning mortgage market, which was built on a governmental subsidy program aimed at improving housing conditions. Due to this program mortgage loans became affordable and by the end of the recent decade the mortgage loans to GDP ratio reached levels similar to those of many developed countries. This also caused housing starts to hit a 15 year high and house prices to overshoot by 2004. House prices and the construction industry also proved to be highly sensitive to the recent financial crisis.

Nevertheless, while the effects of changing mortgage rates or recessions on the housing market are well documented, the discussion on the effects of demographic change is still far from reaching a consensus. In their (in)famous paper, Mankiw and Weil (1989) forecast falling real house prices from the 1990s for the US, based on the aging of the baby boom generation. This forecast led to heated critiques, even before history could convincingly contradict it. Poterba, Weil, and Shiller (1991) could reject demography having a strong effect by investigating the house prices of US cities. DiPasquale and Wheaton (1994) built a macro model to show that the increasing scarcity of land is an important factor in rising house prices, which should offset the effects of aging baby boomers. More recently, Martin (2006) argued that if one takes into account that the real interest rate depends on the age structure as well, then demography does turn out to be the most important factor determining house prices.

The main motivation for analyzing the effects of demographic change on the housing market stems from the fact that in many countries considerable changes are expected to happen (and are already happening) in the age structure. Most notably, in the US the baby boom generation has just started to retire. On a smaller scale, the same is true for the UK and other western countries, and Hungary is also expected to experience a boom in the number of pensioners until 2020. Simultaneously, the share of young adults in the population, who have an increasing demand for housing, is expected to decrease substantially. However, the magnitude of the effects on the housing market of these expected demographical changes is unclear.

The contribution of this paper is twofold: (1) it estimates the price and income elasticities of housing demand on Hungarian data, and (2) it investigates whether aging population can be identified as a demand shifter or not, providing empirical evidence to the debate on the role of demography. I use the baseline housing demand model, which can be found for example in Meen (1990) and Cameron, Muellbauer, and Murphy (2006) and do not address methodological issues arising from using spatial data.

Using a panel data set from 2000 on Hungarian micro-regions I find that the price and income elasticities of demand for houses are both in the region of one third (in absolute value), while aging population has a significant but moderate effect compared to other determinants, like income. The available forecasts on demographic trends and the presented estimation results suggest that the combined effect of shrinking population and changing age structure will, *ceteris paribus*, decrease real house prices at least by 10% by the end of the current decade.

The next section introduces a simple theoretical framework of housing demand. Data is presented in Section 3, while the estimation is carried out in Section 4. Several robustness checks are performed in Section 5. The impact of demographic change is discussed in Section 6. The final section concludes the paper.

2 The simplest model of housing demand

This section briefly derives the inverted demand curve that will be estimated. The baseline model by Meen (1990) and Cameron et al. (2006) is closely followed. Consider a representative agent who maximizes life time utility

$$\int_0^\infty e^{-\rho t} u(t) dt, \tag{2.1}$$

where $u(t) = U(C(t), H(t))$ and $C(t)$ and $H(t)$ denote non-housing composite consumption and the housing stock, respectively. The agent faces the following budget constraint:

$$p_h(t) I_h(t) + C(t) + S(t) = Y(t) + iA(t), \tag{2.2}$$

where $p_h(t)$ is the real price of a dwelling, $I_h(t)$ is the number of new dwellings constructed, $S(t)$ is real net savings, $Y(t)$ is real income, and i is the nominal interest received on net non-housing assets, $A(t)$. The price of the composite consumption good is set to unity. Finally, the law of motion for the housing stock and for non-housing assets are given by

$$\dot{H}(t) = I_h(t) - \delta H(t) \tag{2.3}$$

$$\dot{A}(t) = S(t) - \pi A(t), \tag{2.4}$$

where δ is the amortization rate of the housing stock and π is the rate of inflation (amortization of non-housing savings). It is shown in Appendix A that based on this system one can find the marginal rate of substitution between the housing stock and the consumption good to be (in the deterministic case):

$$\frac{U'_H}{U'_C} = p_h(t)[i + \delta - \pi - \dot{p}_h(t)/p_h(t)]. \tag{2.5}$$

The expression in brackets is known as the user cost of housing. It is positively related to the forgone interest (i) and to the difference of amortization rates ($\delta - \pi$), implying that if houses depreciate faster than alternative assets then holding them is also more expensive. Finally, realizing capital gains ($\dot{p}_h(t)/p_h(t)$) decreases the user cost. One could augment the user cost by taking into consideration further relevant factors like tax rates on interest earnings or the different risk profile of real estate compared to other assets. However, as these factors are neglected in the empirical analysis below, this level of complexity is satisfactory for the current study's purposes.

In the market for the flow of housing services the equilibrium real rental, $r(t)$ is going to clear the market. Therefore, house prices will adjust to clear the asset market, until

$$r(t) = p_h(t)[i + \delta - \pi^e - \dot{p}^e_h(t)/p_h(t)] = p_h(t)UC, \tag{2.6}$$

where e denotes expectations and UC is the user cost. Notice that two assumptions have been made: (1) it is assumed that housing services and the housing stock are proportional, implying that the real rental for housing services may be used as the relevant price and (2) expected inflation and capital gains are no longer substituted with actual values. Equation (2.6) is the heart of most applied work as it states that the unobservable real rental can be proxied by the product of house prices and the user cost of housing.

Now I turn to the second building block of the model. Assume that the quantity of housing services at time t is measured by the per capita housing stock, H/POP_t. Then the demand curve can be written as (like in Cameron et al., 2006):

$$\ln H/POP_t = \alpha \ln Y/POP_t - \beta \ln r_t + \delta z_t, \tag{2.7}$$

where α and β (both assumed to be positive) are the income and price elasticities of housing services, respectively. Additional demand shifters, like demographic trends are captured by z_t. If one substitutes out the real rental by using equation (2.6) and expresses log house prices from the resulting formula then this leads to the following inverted demand relation (making the necessary jump from continuous to discrete time):

$$\ln p_t = 1/\beta \ln H/POP_t + \alpha/\beta \ln Y/POP_t - \ln UC_t + \delta/\beta z_t. \tag{2.8}$$

Before discussing the estimation strategy of equation (2.8) the regional data used in this study is introduced.

3 Data and aggregate dynamics

3.1 Availability of data and descriptive statistics

In applied work which concerns the real estate market, availability of data on house prices usually sets the boundaries of the analysis. In the case of Hungary, real estate transactions are recorded and organized reliably from the end of the 1990s. I use a comprehensive transaction level data set provided by the National Tax and Customs Office (NTC)[1]. The data starts from year 2000 and provides information on the price, size, type and the postal code level location of the real estate.

There are some trade offs to be considered when choosing the unit of analysis in the cross-sectional dimension. Two plausible solutions can be conceived: the county level (NUTS III level, 20 units in the cross-sectional dimension) or the statistical micro-regions (NUTS IV level, 174 units). The main trade offs are the following:

1. Measurement error of house prices (and potentially of other variables) vs. number of cross-sectional units. While on the county level, it is possible to use a hedonic method for estimating house prices, on the micro-regional level, the low number of transactions only allows for using mean or median square meter prices.

2. Homogeneity of the observational unit vs. units being distinct submarkets. Statistical micro-regions were formed with the aim of measuring detailed regional development and they can be considered to be homogeneous. In contrast, counties are highly heterogeneous geographically and socially. Nevertheless, counties do have the advantage of functioning better as distinct housing submarkets. This is crucial as estimation

[1] Nemzeti Adó- és Vámhivatal.

techniques used here implicitly assume this property.

3. Data availability in the time dimension. Currently, per capita income is only available until 2008 at the micro-regional level, while other proxies of income as net wages are available until 2010 at the county level.

The trade offs presented above could motivate either administrative level. Therefore, the detailed results of the micro-regional level are presented and the county level is left as a robustness check.

Table 3.1 shows some descriptive statistics for the pooled micro-regional data. First, due to the uneven distribution of transactions, micro-regional house prices are measured with high idiosyncratic error.[2] This will also make it necessary to weight observations in the regression analysis.[3] Second, the stock of dwellings increased considerably. If one takes a closer look at the data, the micro-regions which experienced above 4% growth rate in any year, were all concentrated in the agglomeration of Budapest.[4] This is in line with Kovács and Dövényi (2006), who discuss the development of the "metropolitan periphery" in detail. Finally, the percentage change in per capita dwellings indicates that in most micro-regions the per capita housing stock is upward trending.

The county level statistics are presented in Table 3.2. Contrary to the micro-regional data that ends in 2008, county level data is updated until 2010. The price for these additional two years is that per capita income is proxied by per capita net wages, which is arguably a noisy proxy of per capita income.

[2] In four occasions, after eliminating unusable records, there were no transactions left for the given micro-region and year.

[3] About one fifth of the housing stock and over one third of the transactions can be found in Budapest, which is represented by a single micro-region. A straightforward way to lower the concentration (and also to increase the number of cross-sectional units) would have been introducing each district of the capital independently, however, data availability prevented this.

[4] The single exception is the Hévízi micro-region, which is known for its excellent location and thermal lake.

Table 3.1: Descriptive statistics of micro-regional pooled data, 2000-2008.

Variable	Median	Mean	Std. dev.	Min	Max	Source
Real m^2 price	621	669	255	196	1614	NTC
% change in m^2 price	4.7	5.8	14.8	-59.6	118.9	NTC
Frequency of transactions	106	439	2431	0	47954	NTC
Real per capita income	3003	3142	947	1258	6651	HRD
Stock of dwellings	14676	23931	64538	3235	881000	HCSO
% change in stock of dwellings	0.37	0.48	0.96	-4.0	8.3	HCSO
% change in per capita dwellings	0.91	0.79	0.89	-4.1	4.9	HCSO

Note: Real m^2 price and real per capita income are in 2010 forints, converted to US dollars at 200 HUF/USD. NTC - National Tax and Customs Office; HRD - Hungarian Regional Database; HCSO - Hungarian Central Statistical Office. Please consult Appendix B for detailed definitions of variables.

As expected, the variation of price changes is smaller compared to the previous case, which is the result of aggregation and smaller measurement error. As there are no missing values for prices, this is a balanced panel. Square meter prices are obtained from hedonic regressions, which means that composition effects due to the size, type and location of the real estate were controlled for. Note that by adding the years 2009-2010 median and mean changes of real house prices decreased dramatically due to the recent financial crisis. In fact, real house prices in 2010 were not far from their year 2000 value and, compared to 2008 they dropped by an average of 18%.

Out of 200 observations in 24 cases the percentage change in per capita dwellings was negative, but the majority of this is due to 2001/2000, when the population of Budapest decreased by over 50 thousand, assumably due to migration.[5] This resulted in population

[5] As there was a census in 2001 the case might be that data before 2001 were not adequately revised. However, as no evidence was found in support of this suspicion, I assumed the data to be correct.

Table 3.2: Descriptive statistics of county level pooled data, 2000-2010.

Variable	Median	Mean	Std. dev.	Min	Max
Real m^2 price	728	764	230	375	1609
% change in m^2 price	-0.15	1.5	8.9	-17	29
Frequency of transactions	2019	3420	6115	120	47961
Real net monthly wage	562	557	82	372	846
Stock of dwellings	165217	208881	164340	87829	889757
% change in stock of dwellings	0.5	0.58	0.38	-0.39	2.3
% change in per capita dwellings	0.91	0.70	1	-2.9	3.0

Note: Real net monthly wage is from HCSO. Please consult the notes of Table 3.1 and Appendix B for sources of variables and detailed definitions.

increasing faster than the house stock outside the capital. Otherwise, per capita dwellings are steadily upward trending at the county level.

Finally, variation in the share of young adults between and within micro-regions would be welcome, as this variable will be one focus of the regression analysis. Figure 3.1 shows how the share of young adults changed from 2000 to 2008. At the country level their share decreased from 15.7% to 13.9%.[6] Since real house prices are also decreasing since 2004 at the country level, there is a valid concern that the estimates will be spurious to some extent. However, as in many micro-regions the share of young adults stagnated or even increased opposed to the country wide trend, this should provide sufficient variation in order to capture the true effect of this variable.

[6] One striking fact about this map is that Central Hungary suffered the biggest decrease in the share of young adults. While arguably this region offers the best employment opportunities, the cost of living (especially real estate) is also here the most expensive within the country. Still, factors like international migration and lower initial fertility rate probably also contributed to this process.

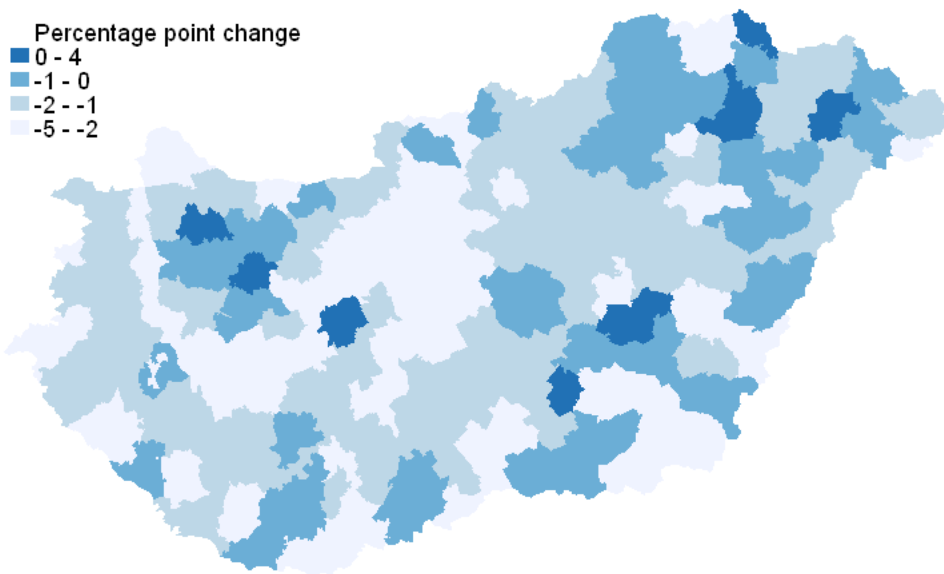

Figure 3.1: Change of the share of population aged 20-29 from 2000 to 2008. Source of data: HCSO.

3.2 Aggregate dynamics

Many countries experienced housing bubbles in the past decade.[7] As the data here covers the rather short period from 2000, it would discredit the results if there was also a real estate bubble during this period in Hungary. Though the presence of a bubble is not tested formally, I argue, based on aggregate dynamics, that it is unlikely that there was a real estate bubble.

The introduction of the governmental subsidy program revitalized mortgage lending by considerably lowering the effective mortgage rate in 2002.[8] Following the previous section, it is clear that if interest rates decrease then the user cost also decreases. Poterba (1984) extensively analyzed the consequences of a reduction in the user cost in a simple theoretical

[7] For a recent overview on the topic see Malkiel (2010).

[8] According to Kiss and Vadas (2007) the effective nominal mortgage rate was above 12% in 2001 and came down to 6% by the second quarter of 2002. Inflation also decreased from 9% in 2001 to 5% in 2002 and 2003. As the regulations of the subsidy program tightened in December 2003, foreign currency based loans appeared in the market, often offering even lower mortgage rates than the subsidezed rate. Until the beginning of the recent crisis the mortgage rate remained around 6%.

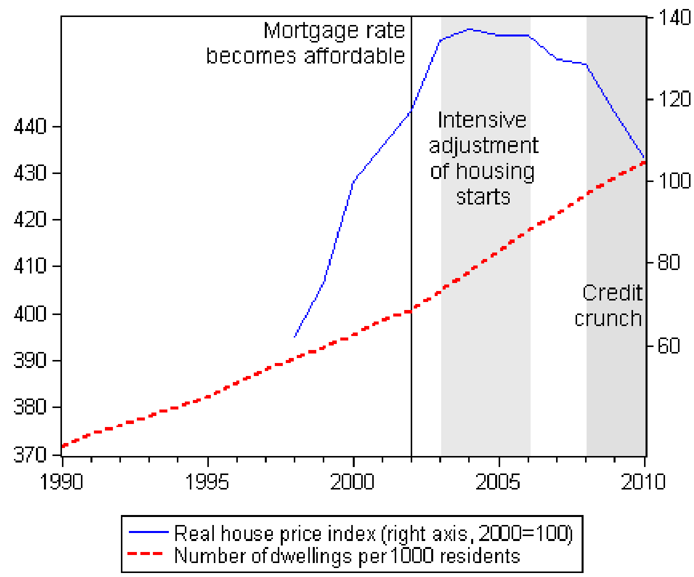

Figure 3.2: Recent major events and episodes in the Hungarian housing market. Sources: own calculations based on data from HCSO and NTC.

framework. According to the well-known argument, if there is a negative user cost shock, then the sluggish adjustment of quantity causes the price to overshoot as it reaches the new saddle path. Afterward prices gradually decrease, until the quantity reaches its new steady state.

Aggregate real price and quantity are plotted in Figure 3.2. The events are consistent with those suggested by Poterba (1984). After the introduction of the subsidy program, prices began to overshoot. As quantity started to adjust rapidly by 2004, prices gradually declined presumably because they started to converge to the new steady state. Due to the crisis, falling income and rising interest rates (opposed to the US, where mortgage rates were also declining) caused price depreciation to accelerate from 2009. Therefore, I conclude that aggregate price dynamics were driven by fundamentals and strong evidence could not be found for the presence of a real estate bubble.

Figure 3.3 shows the projected path of the share of young adults in the population,

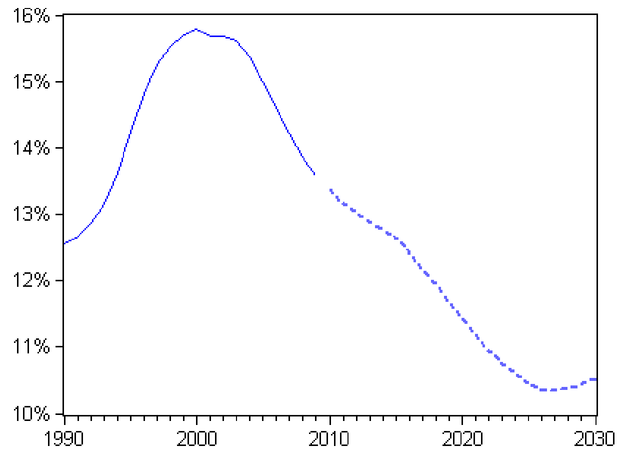

Figure 3.3: Share of population aged 20 to 29 and its projected path. Source of data: HCSO - Demographic Research Institute.

which is the primary focus of the paper. According to the forecast, until 2020 their share will decrease by about 2 percentage points. If my estimation results in the coming sections are correct, then they imply that the steep increase in house prices before 2000 were partly driven by the increasing share of young adults, further supporting the claim, that fundamentals drove house prices in Hungary.

4 Estimation strategy and results

In this section the following reduced-form model is estimated, which is motivated by equation (2.8):

$$p_{it} = \rho p_{i,t-1} + \gamma_1 (y/pop)_{it} + \gamma_2 (h/pop)_{it} + \gamma_3 UC_t + \gamma_3 z_{it} + c_i + u_{it}, \tag{4.1}$$

where $\gamma_1 (= \alpha/\beta)$ is the income elasticity of house *prices* and $\gamma_2 (= -1/\beta)$ is the inverse of the price elasticity of housing demand. Recall that α and β are the income and price elasticities of housing demand, respectively. Subscript i stands for micro-region ($i = 1 \dots N$, $N = 174$) and t for year ($t = 2000 \dots 2008$). Lower case letters indicate natural logarithms of variables. The user cost will enter the regression in levels, as in several cases it is negative.[1] All variables are weighted by the square root of the micro-regional housing stock of year 2000, i.e. multiplied by $\sqrt{h_{i,t=2000}}$.

It is assumed that c_i are fixed regional effects allowed to be correlated in an arbitrary way with other right hand side variables. For example, labor market factors are likely to be present in c_i, as labor market conditions are probably the main reason why people move and, therefore, enter the real estate market. Of course, labor market factors are likely to be correlated with per capita income and with the per capita housing stock.

Estimating equation (4.1) using the standard fixed effects (within) or first difference (FD) estimator is problematic at best. There are several sources of biases:

1. h/pop is likely to be determined simultaneously with house prices because of supply effects.

[1] For more detailed definitions of the variables consult Appendix B

2. It is common to include the lagged dependent variable in the right hand side to capture sluggish adjustment. However, using the within or the first difference estimator will lead to bias in this case.

3. Measurement error in the dependent variable.

4. The user cost also includes the lagged dependent variable for capturing expectations, causing the same problem as in point 2.

5. The error term, u_{it}, is likely to be serially correlated.

6. The housing stock is only a proxy for housing services as it is not adjusted by size or quality, leading to measurement error in h/pop.

The errors could also be spatially correlated, leading to a non block-diagonal error covariance matrix, causing clustered errors to be biased and inconsistent. However, I will not address spatial issues in this paper, and assume that they are only second order in importance compared to the listed concerns. From the potential sources of biases it is argued that the endogeneity of h/pop is the most severe from the perspective of estimating the price and income elasticities of housing demand.

4.1 Within or FD estimator?

It is likely that house prices are measured with the biggest error among all the variables (see Table 3.1). This is due to the insufficient number of transactions from which average square meter prices are calculated. As the stock of dwellings is highly correlated with the number of observations between micro-regions, weighting the observations by the stock of dwellings should somewhat reduce the problem caused by measurement error. Hamermesh (1989) argues that first differencing the data with measurement error in the dependent variable usually leads to imprecise estimates. So the presence of measurement error could support the within estimator.

The within estimator should also be preferred when the lagged dependent variable is not instrumented. For an AR(1) model, Hsiao (2003, pp. 71-72) shows that under the assumptions of weakly dependent variables and $|\rho| < 1$, the bias of the within estimator is of order T^{-1} while the FD estimator's bias does not decrease as $T \to \infty$. Based on Kiss and Vadas (2007) a good argument can be made for $|\rho| < 1$, but all variables are highly persistent, so the within estimator should not have an advantage in this respect, especially after instrumenting the lagged dependent variable.

Furthermore, strict exogeneity, which is assumed if one uses the within estimator is probably too restrictive. Under strict exogeneity u_{it} should be uncorrelated not only with current but with all past and future values of the independent variables as well, because time averages enter the condition $E(u_{it} - \bar{u}_i | x_{it} - \bar{x}_i) = 0$, where x_{it} are right hand side variables. A less restrictive assumption is sequential exogeneity, as it only requires the error term to be uncorrelated with the present and all past values of the right hand side variables.

The assumption of strict exogeneity could easily be violated in (4.1). Suppose a bypass road is opened to transportation, reducing the negative externality of transit traffic in a city. This should cause a positive shock to house prices by making the city more attractive to higher income households ($Corr(u_{it}, (y/pop)_{i,t+1}) > 0$). Moreover, if households' valuation of transit traffic is correlated with household size (e.g. households with children value the clean environment more) then this should lead to $Corr(u_{it}, (h/pop)_{i,t+1}) < 0$. If such hypotheses seem credible then the strict exogeneity assumption should not hold and, therefore, the within estimator should certainly not be preferred.

Finally, serial correlation in u_{it} could turn out to be an important factor when it comes to the decision between the within or the FD estimator. I implement the test for serial correlation for the FD estimator as described in Wooldridge (2010, pp. 319-320). If the differenced errors were uncorrelated that would imply that the FD estimator is probably more efficient. Estimates of first order autocorrelation of the differenced errors are provided in Table 4.1. The coefficients of serial correlation turn out to be negative in all specifications but are close to zero. Therefore, the FD estimator should be preferred in this respect. Based on this argumentation a stronger case can be made for the FD estimator.

4.2 Estimation results

Table 4.1 shows the results of estimating (4.1) in six different specifications, using the FD estimator.[2] In the IV(1) specification $(h/pop)_{i,t}$ is instrumented by $(h/pop)_{i,t-3}$, and the three year lagged construction cost, $(ccost_{i,t-3})$. The rationale for using three year lagged variables is that the design and construction period together for dwellings usually consumes less than three years.[3] Therefore, while per capita housing stock is highly persistent, implying $Corr((h/pop)_{i,t}, (h/pop)_{i,t-3}) >> 0$, it should not be correlated with house prices that are distant in the future, because in the short run supply is highly inelastic.

As Meen (1996) argues, due to the low price elasticity of new housing starts there should not be any significant relationship between construction costs and house price shocks, as the latter are typically driven by demand factors, implying $Corr(u_{i,t}, ccost_{i,t-3}) = 0$. But *ceteris paribus* higher construction costs should lead to fewer building starts, and to a smaller stock in the future, so $Corr((h/pop)_{i,t}, ccost_{i,t-3}) < 0$ is expected. Therefore, proxies for construction costs should also be valid instruments. I use labor costs in the

[2] I carry out the estimation using the algorithm provided by Schaffer (2005).

[3] The Hansen-J statistic for overidentification is also in line with this argument. When h/pop is instrumented by one or two year lagged h/pop and $ccost$, the null hypothesis of valid instruments has a p-value of below 5%, while with three year lagged variables, the same p-value is well above 5%.

construction industry to proxy for construction costs.[4]

The expected region of the estimated parameter on h/pop is below minus one as this is the inverse of the price elasticity of housing demand. As shown in Table 4.1, the estimated coefficient on h/pop is not significantly different from one at 5% when estimated by least squares, regardless of controlling for time fixed effects. However, after instrumenting h/pop in IV(1), the coefficient becomes significantly smaller than minus one. The direction of the bias of specifications (1) and (2) is consistent with entrepreneurs timing the completion of new dwellings to maximize profits. Consider the case of the development of infrastructure, like developing the sewage system. During the construction, which is typically measured in years, house prices are not likely to fully adjust due to the costs and negative externalities of construction. However, by the time the project is finished, prices should adjust and the supply of houses should also increase. Mechanisms like this should cause the estimated coefficient on h/pop, to be biased upwards. And in fact, the results of specifications (1)-(2) and IV(1) are consistent with this story.

In the IV(1)-IV(3) specifications the FD estimator finds the implied price elasticity of housing demand to be in the range of -0.22 to -0.30. However, as housing services are proxied by the per capita housing stock, these elasticities only capture extensive margin effects. Therefore, it is argued that the true elasticities are greater in absolute value, because of substantial fixed costs. Consider a family who is renting a flat and the real rentals decline in their neighborhood. Assume that they have strong preferences for their neighborhood but they would prefer a dwelling with an extra bedroom. Although they have the option to rent an additional flat, for many households it would seem more rational to rent a bigger flat instead.

[4] The first stage results for instrumenting the per capita housing stock is presented in Table C.1. Both instruments are significant and have the expected sign.

Table 4.1: Estimation results for the micro-regional panel using the FD estimator.

First-difference estimation						
Dependent variable: micro-regional log real house price ($p_{i,t}$)						
	(1)	(2)	IV(1)	IV(2)	IV(3)	IV(4)
$p_{i,t-1}$	-0.113	-0.353	-0.327	-0.282	-0.271	-0.179
	[0.051]	[0.045]	[0.055]	[0.222]	[0.149]	[0.134]
$(y/pop)_{it}$	0.909	0.546	0.450	Restr.	Restr.	Restr.
	[0.126]	[0.256]	[0.328]	to 1	to 1	to 1
$(h/pop)_{it}$	-0.550	0.108	-4.627	-3.398	-3.352	-1.561
	[0.470]	[0.535]	[1.204]	[1.278]	[1.088]	[0.807]
$pop20-29_{it}$						0.075
						[0.031]
UC_{it}	-0.274	-0.212	-0.192	-0.180	-0.167	-0.127
	[0.036]	[0.047]	[0.051]	[0.110]	[0.183]	[0.188]
Year fixed effects	No	Yes	Yes	Yes	Yes	Yes
Obs.	1034	1034	861	861	861	861
R-squared	0.148	0.394	0.049	0.281	0.281	0.296
Regions	173	173	173	173	173	173
1st order serial corr.	-0.109	-0.016	-0.005	-0.052	-0.060	-0.158
	[0.045]	[0.046]	[0.031]	[0.039]	[0.039]	[0.042]
Hansen-J p-value			0.65	0.63	0.63	0.38

Note: standard errors in brackets are based on the clustered covariance matrix by micro-regions. All variables are differenced (including year fixed effects and instruments) and weighted by the square root of the stock of dwellings. y/pop - log per capita permanent income, h/pop - log per capita stock of dwellings, UC - user cost, $pop20-29$ - percent of population aged between 20 and 29. In IV(1) $(h/pop)_{it}$ assumed to be endogenous (instruments: $(h/pop)_{i,t-3}$, $ccost_{i,t-3}$), in IV(2) h/pop_{it} and $p_{i,t-1}$ assumed to be endogenous (instruments: $(h/pop)_{i,t-3}$, $ccost_{i,t-3}$, $p_{i,t-2}$), in IV(3) $(h/pop)_{it}$, UC_{it} and $p_{i,t-1}$ assumed endogenous (instruments: $(h/pop)_{i,t-3}$, $ccost_{i,t-3}$, $p_{i,t-2}$, UC_{t-1}) and in IV(4) $(h/pop)_{it}$, UC_{it}, $p_{i,t-1}$ and $pop20-29_{i,t}$ assumed endogenous (instruments: $(h/pop)_{i,t-3}$, $ccost_{i,t-3}$, $p_{i,t-2}$, UC_{t-1} and $pop20-29_{i,t-1}$).

Further specifications impose the restriction that the income elasticity of house *prices* (γ_1) is equal to one. This restriction is motivated by three reasons: (1) in the first specification it is consistent with the data, (2) year fixed effects seem to catch the effects of changes in income and (3) Kiss and Vadas (2007) also find this long run elasticity to be one, using county level quarterly panel data from 1997 to 2002. As $\gamma_1 = \alpha/\beta$, this restriction implies that the price and income elasticities will be restricted to be equal in absolute value.

The specification IV(2) addresses the issue of endogeneity caused by the lagged dependent variable, instrumenting it by $p_{i,t-2}$. Note that h/pop is still instrumented as in IV(1). The coefficients found by Kiss and Vadas (2007) on 1 and 2 quarter-lagged prices in different specifications imply a value for ρ to be in the range of 0-0.3 in yearly data. However, specifications in Table 4.1 find negative and mostly insignificant coefficients.

As described in Appendix B, the user cost captures expected capital gains with lagged house price changes. Therefore, the user cost is partly driven by the lagged dependent variable, which causes bias if estimated by the within estimator or by first differencing. The IV(3) specification addresses this issue by instrumenting UC_{it} with $UC_{i,t-1}$. Though the estimated coefficient on UC_{it} does not change dramatically, its standard deviation does increase, which is likely to be a sign of weak instruments.

From equation (2.8) one would expect $\hat{\gamma}_3$ (the coefficient on the user cost) to be larger. In fact, the estimates in all specifications are in the region of -0.2, which might be an indication of the user cost being misspecified.[5] As noted in Section 3.2, the single, large change in the mortgage rate took place in 2002 and possibly led prices to overshoot, implying delayed effects. But these effects are probably captured by the year fixed effects. As the user cost is not the main focus of the paper, it is not discussed further.[6]

Finally, the last specification augments IV(3) with $pop20-29$, which is the fraction of the micro-regional population aged 20 to 29. The motivation for adding this variable to the right hand side is based on Mankiw and Weil (1989), who show that this is the age

[5] The estimate of -0.2 implies that if the user cost increases by 1 percentage point then house prices decrease by 0.2%

[6] If the reader is interested in the mortgage rate elasticity of house prices, consult Kiss and Vadas (2007).

cohort who have increasing demand for real estate with aging.[7] Therefore, the hypothesis of this paper is that their share in the population could be a possible demand shifter. The exogeneity of this variable depends on the validity of the assumption about regional fixed effects. If the regional fixed effects are indeed fixed throughout the sample period and capture factors, which induce people to move, then assuming exogeneity of young adults is probably acceptable. However, if for example labor market conditions changed rapidly between and within micro-regions then the exogeneity assumption of $pop20-29$ is questionable, as this cohort is the most responsive in terms of mobility.[8] I instrument $pop20-29$ by its own lagged value, allowing for sequential exogeneity.

The share of young adults enters significantly in the FD regression, and leads the coefficient on h/pop to increase considerably. Untabulated calculations show that while h/pop is negatively correlated with $pop20-29$ within micro-regions (as a result of upward trending h/pop and downward trending $pop20-29$), they are positively correlated between regions (young households are overrepresented in regions with relatively small household size, like the capital city). However, the former, negative correlation is not entirely spurious. For example, Lindh and Malmberg (2008) shows on Swedish data that demographic composition is closely related to construction activity, i.e. typically young/middle aged adults build houses. Therefore, when one is interested in the income and price elasticity of housing demand, then the share of young adults should not be included in the regression.

The estimated coefficient on young adults is positive and significantly different from zero. It implies that if *ceteris paribus* the share of young adults increases by 1 percentage

7 Note that Mankiw and Weil (1989) do not use this variable, as their data enables them to estimate the demand of each age cohort. However, using the share in the population of a distinguished age cohort to proxy for demand effects is not novel in the literature. For example, Geanakoplos, Magill, and Quinzii (2004) use the share of the middle-aged cohort in explaining stock market returns. Choosing the 20-29 aged cohort is supported by the fact that according to Iacovou and Skew (2010) half of the women (men) aged above 25.0 (27.6) are already living away from their parental home in Hungary.

8 For example, according to the HCSO, in 2005 28% of internal migration was due to this cohort, while their share in the population was 15%.

point than house prices will increase by 7.2%. Further specifications discussed in the next section suggest that the effect is actually more modest.

5 Robustness checks

It is crucial for the estimation of the price and income elasticities of housing demand that the endogeneity of the per capita housing stock is addressed adequately. Recall that it was instrumented by its own 3 year lagged value and by 3 year lagged construction costs. As a robustness check, I estimate the model using the within estimator and replacing the instrument $(h/pop)_{i,t-3}$ with a deterministic trend. A deterministic trend should capture the effects of variables that are correlated with long term structural changes, for example which cause average household size to decrease.[1]

Table 5.1 shows similar specifications as in Table 4.1. The only differences are the following: (1) it is estimated by the within estimator, (2) the instrument $(h/pop)_{i,t-3}$ is replaced by a deterministic trend and (3) year fixed effects are only included for 2007 and 2008.[2] The estimated coefficients on $(h/pop)_{i,t}$ are statistically and economically identical with those in Table 4.1 for specifications IV(2)-IV(4). The results of the first stage for instrumenting the per capita housing stock can be found in Table C.2. While the deterministic trend enters the first stage with a highly significant coefficient, the lagged construction cost becomes insignificant when estimated by the within estimator.

As discussed in Section 3, a priori the optimal size of the cross sectional unit is not trivial. Therefore, I estimate equation (4.1) at the county level in first differences and present the results in Table 5.2. First, the estimated coefficients are quite close to those estimated on micro-regional data. The implied elasticity of housing demand is somewhat larger, between -0.33 to -0.4. The share of young adults enters the regression significantly, but its estimated coefficient is smaller and falls in the region of 3-4%, opposed to the

[1] Meen (1990) also uses a deterministic trend as an instrument for per capita housing stock. Obviously, using the FD estimator with a trend as an instrument is not feasible.

[2] Due to the trend all year fixed effects cannot be included.

Table 5.1: Estimation results for the micro-regional panel using the within estimator.

Within estimation						
Dependent variable: micro-regional log real house price (p_{it})						
	(1)	(2)	IV(1)	IV(2)	IV(3)	IV(4)
$p_{i,t-1}$	0.390	0.250	0.193	0.254	0.244	0.166
	[0.038]	[0.048]	[0.057]	[0.090]	[0.087]	[0.084]
$(y/pop)_{it}$	0.372	0.780	0.667			
	[0.076]	[0.129]	[0.215]			
$(h/pop)_{it}$	-0.920	-0.597	-1.971	-3.523	-3.193	-1.760
	[0.234]	[0.252]	[0.645]	[0.891]	[1.089]	[0.823]
UC_t	-0.272	-0.354	-0.121	-0.085	-0.155	-0.198
	[0.051]	[0.058]	[0.068]	[0.093]	[0.280]	[0.269]
$pop20-29_{it}$						0.032
						[0.014]
y07		-0.080	-0.047	-0.051	-0.055	-0.055
		[0.009]	[0.012]	[0.012]	[0.014]	[0.014]
y08		-0.085	-0.047	-0.048	-0.053	-0.054
		[0.013]	[0.016]	[0.021]	[0.017]	[0.016]
Obs.	1209	1209	1034	1034	1034	1034
R-squared	0.380	0.433	0.113	0.527	0.540	0.590
Regions	174	174	173	173	173	173
Hansen-J p-value			0.50	0.48	0.36	0.97

Note: standard errors in brackets are based on the clustered covariance matrix by micro-regions. All variables are weighted by the square root of the stock of dwellings. y/pop - log per capita permanent income, h/pop - log per capita stock of dwellings, UC - user cost, $pop20-29$ - percent of population aged between 20 and 29. In IV(1) $(h/pop)_{it}$ assumed to be endogenous (instruments: det. trend, $ccost_{i,t-3}$), in IV(2) h/pop_{it} and $p_{i,t-1}$ assumed to be endogenous (instruments: det. trend ,$ccost_{i,t-3}$, $p_{i,t-2}$), in IV(3) $(h/pop)_{it}$, UC_{it} and $p_{i,t-1}$ assumed endogenous (instruments: det. trend, $ccost_{i,t-3}$, $p_{i,t-2}$, UC_{t-1}) and in IV(4) $(h/pop)_{it}$, UC_{it}, $p_{i,t-1}$ and $pop20-29_{i,t}$ assumed endogenous (instruments: det. trend, $ccost_{i,t-3}$, $p_{i,t-2}$, UC_{t-1} and $pop20-29_{i,t-1}$).

baseline estimate of 7%. The result is unchanged after $pop20-29_{jt}$ is instrumented by its lag, as in specification IV(5). This implies that the results are robust under the less restrictive assumption of sequential exogeneity.

Second, recall that county level data is updated until 2010. Thus, the recent crisis is captured by this regression. Unfortunately, the frequency of the time dimension is probably still too low to gain robust estimates on the user cost, which is simply proxied here by the mortgage rate. In 2009 and 2010 the mortgage rate increased sharply with declining real house prices, which leads one to expect large and negative coefficients on the mortgage rate, like in specifications (1) and IV(1'). However, after restricting the coefficient on permanent income, estimated coefficients on *mrate* become similar to those in Table 4.1.

To address the issue of the robustness of the results with respect to the definition of permanent income, I estimate IV(3) of Table 4.1 and IV(4') of Table 5.2 under different definitions of permanent income and report the results in Table 5.3. So far, permanent income was defined by $(Y/POP)_{it} = 0.5INC_{it} + 0.3INC_{i,t-1} + 0.2INC_{i,t-2}$.[3] The definitions differ in the weights that are attributed to the current and past 2 years of income but the sum of weights is always restricted to unity. The main result of Table 5.3 is that while the micro-regional model is robust to the definition of permanent income, the county level estimates are quite sensitive. Besides the obvious reason of fewer observations, this might also be caused by that income is measured by a weaker proxy at the county level, as discussed in Appendix B.

The regression outputs in Tables 4.1 and 5.2 weight all variables by the stock of dwellings in 2000. The reason for using the stock of a single year is that this way weights do not vary within cross sectional units, guaranteeing each year to have the same weight. However,

[3] An early example using a similar definition can be found in Carliner (1973).

Table 5.2: Estimation results for the county level panel using the FD estimator.

First-difference estimation						
Dependent variable: county level log real house price ($p_{j,t}$)						
	(1)	IV(1')	IV(2')	IV(3')	IV(4')	IV(5)
$p_{j,t-1}$	0.386	0.367	0.258	0.329	0.030	0.025
	[0.058]	[0.061]	[0.115]	[0.116]	[0.064]	[0.067]
$(y/pop)_{jt}$	0.663	0.442	0.604	0.334	Restr.	Restr.
	[0.120]	[0.134]	[0.185]	[0.208]	to 1	to 1
$(h/pop)_{jt}$	-2.320	-3.001	-3.272	-2.577	-2.774	-2.514
	[0.225]	[0.300]	[0.270]	[0.367]	[0.488]	[0.810]
$pop20-29_{jt}$					0.032	0.041
					[0.012]	[0.021]
$crisis_t$				-0.038	-0.003	-0.002
				[0.011]	[0.008]	[0.008]
$mrate_t$	-1.132	-0.888	-0.648	-0.503	-0.135	-0.164
	[0.225]	[0.243]	[0.198]	[0.200]	[0.165]	[0.183]
Obs.	160	140	140	140	140	140
R-squared	0.718	0.690	0.679	0.713	0.484	0.486
Counties	20	20	20	20	20	20
Serial Corr.	-0.108	-0.164	-0.002	-0.099	0.138	0.135
Test	[0.077]	[0.074]	[0.078]	[0.076]	[0.081]	[0.081]
Hansen-J		0.09	0.14	0.25	0.12	0.13
p-value						

Note: standard errors in brackets are based on the clustered covariance matrix by counties. All variables are differenced (including year fixed effects and instruments) and weighted by the square root of the stock of dwellings. y/pop - log per capita permanent income, h/pop - log per capita stock of dwellings, $mrate$ - mortgage rate, $pop20-29$ - percent of population aged between 20 and 29, $crisis$ - 1 if $t = 2009$ or $t = 2010$, zero otherwise. In IV(1') $(h/pop)_{jt}$ assumed to be endogenous (instruments: $(h/pop)_{j,t-3}$, $ccost_{j,t-3}$), in IV(2')-IV(4') h/pop_{jt} and $p_{j,t-1}$ assumed to be endogenous (instruments: $(h/pop)_{j,t-3}$, $ccost_{j,t-3}$, $p_{j,t-2}$) and in IV(5) $(h/pop)_{it}$, $pop20-29_{jt}$ and $p_{i,t-1}$ assumed endogenous (instruments: $(h/pop)_{j,t-3}$, $ccost_{j,t-3}$, $p_{j,t-2}$, $pop20-29_{j,t-1}$).

Table 5.3: Robustness of results to different definitions of permanent income.

Estimated coefficients under different definitions of permanent income from IV(3) of Table 4.1 (IV(4') of Table 5.2).					
Weight of income t in permanent income:	0.5	0.4	0.5	0.7	1
Weight of income $t-1$ in permanent income:	0.3	0.5	0.5	0.3	0
Weight of income $t-2$ in permanent income:	0.2	0.1	0	0	0
Estimated coefficient on:					
h/pop	**-3.35**(-2.8)	**-3.26**(-1.57)	**-3.13**(-1.40)	**-3.07**(-2.40)	**-2.98**(-3.90)
$pop20-29$	∅ (**0.03**)	∅ (**0.05**)	∅ (**0.04**)	∅ (0.02)	∅(-0.01)
$UC/mrate$	-0.17(-0.14)	-0.14(**-0.37**)	-0.17(**-0.28**)	-0.18(0.04)	-0.18(**0.49**)
Note: bold estimated coefficients indicate statistical significance at 5%.					

other potential candidates for weights are population and the number of transactions.[4] I present the results under different weighting schemes in Appendix C for the FD estimator using both the micro-regional and the county panel.

As transactions are much more concentrated than either population or the stock of dwellings,[5] bigger deviations from the baseline results are expected in the case of using transactions as weights. In fact, weighting the variables with the number of transactions leads to smaller implied price elasticities. But this deviation is only present in the micro-regional estimates. The coefficients on the share of young adults remain the same within the micro-regional data (around 7%) and the county level data (3-4%).

Counties form better distinct house markets than micro-regions do. This is important because young adults are the most mobile in the population, implying that they can respond to many shocks in u_{it} by moving. Though I instrument $pop20-29_{jt}$ with its own lagged value, concerns still might be valid about endogeneity, especially for the micro-regional data. Therefore, I prefer the estimates on this variable obtained from the county data.

4 For example, Polinsky and Ellwood (1979) use the number of observations as weights when estimating inverted demand equations from grouped data.

5 Over 30% of transactions take place in Budapest, while 17% of the population lives in Budapest and 20% of the housing stock can be found in the capital.

6 How big is the implied effect of demographic change?

The presented specifications indicate that if *ceteris paribus* the share of young adults increases by 1 percentage point in the total population, then house prices will increase by 3-4%. As the share of young adults can be forecasted with little uncertainty, it is straightforward to calculate the long run effect of aging.[1]

As already shown in Figure 3.3, the share of young adults is projected to decrease by 2 percentage points by the end of the current decade (from 2010). This implies that house prices should decrease by about 6-8% due to aging.

Besides the composition effect, there is also the more obvious effect of the shrinking population. According to the HCSO - Demographic Research Institute the population will decrease by 1.7% until 2020. If one assumes a price elasticity of negative one half (which is reasonable after controlling for young adults) then this effect should further decrease real house prices by 3.4%. Thus, even if one approaches the results with some caution, the estimated total effect of demographic trends on house prices cumulate to -10% by 2020.

However, this estimated effect is moderate compared to the potential effects of other determinants. For example, based on recent forecasts made by the Central Bank of Hungary,[2] GDP growth will be 2.9% and 3% in 2011 and 2012, respectively. Therefore, if one assumes that per capita GDP will grow by 3% annually then until 2020 real house prices should increase by 34% due to economic growth.

[1] In the speculation carried out in this section I do not take into accout the possible effects of international migration. In principle, this should not have a major effect on the results, as according to statistics from Eurostat, only 0.2% of the population were foreigners in 2010 in Hungary, though their number is steadily increasing since 2003.

[2] Central Bank of Hungary (2011)

Though aggregate effects might turn out to be modest, individual effects could still be substantial. Song (2010) demonstrates with a DSGE model that credit constrained households are highly responsive to house price changes. More importantly, the results derived by Li and Yao (2007) from a life-cycle model imply that young homeowners and renters are the biggest winners if house prices depreciate as their future housing consumption costs decrease. In contrast, for old homeowners this is clearly not the case, as large part of their savings (what they intend to consume or leave as a bequest) are accumulated in housing wealth, thus decreasing real house prices might compel them to reduce their consumption.[3]

Unfortunately, the empirical literature could not yet reach a consensus on this matter. Based on data from the UK, Campbell and Cocco (2007) found that the house price elasticity of consumption for old homeowners is 1.7, while for the young ones it is insignificant. Attanasio et al. (2009) argue that Campbell and Cocco (2007) use a misspecified model and reach a conclusion that is less intuitive: the same elasticity is highest for the young, but only high as 0.2. Clearly, methodological issues have to be resolved, before a final verdict can be reached, but old homeowners may turn out to be the biggest losers of current demographic trends.

[3] It is natural to compare this welfare effect to the increasing financial burden on the young caused by aging through the public pension system. See Cerny, Miles, and Schmidt (2010) on the topic of aging, pension systems and the housing market.

7 Conclusion

This study has used a simple model of housing demand and brought it to Hungarian regional data. Estimation results and available forecasts on demographic trends suggest that real house prices will, *ceteris paribus*, decrease at least by 10% due to the shrinking and aging population in Hungary by 2020.

Given the volatility of house prices and their responsiveness to factors like income or interest rates, the estimated aggregate effect of notable demographic change can be best labeled as moderate. Still, certain groups of the population might be sensitive even to such an effect. For example, if older homeowners expect capital gains from their dwellings to be similar to those experienced at the beginning of this century, then they might be compelled to make painful adjustments in their consumption plans.

The income and price elasticities of housing demand are found to be highly inelastic, both estimated to be one third (in absolute value). The true elasticities are arguably larger, but better proxies were not found for housing services than the per capita housing stock.

The results presented in this paper are far from conclusive. Most importantly, external validity should be checked by using data from different countries. As I used the baseline model of housing demand and tackled methodological issues with a bias toward using simplistic solutions, many improvements could be made. The adoption of estimation techniques used by Cameron et al. (2006), who address spatial issues arising from using regional panel data are certainly among these. Nevertheless, one should certainly take away the idea, that major demographic changes on the way will have a non negligible impact on the housing market.

References

Attanasio, O. P., L. Blow, R. Hamilton, and A. Leicester (2009). "Booms and Busts: Consumption, House Prices and Expectations". *Economica 76*(301), 20-50.

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.

Campbell, J. Y. and J. F. Cocco (2007). "How do House Prices Affect Consumption? Evidence from Micro Data". *Journal of Monetary Economics 54*(3), 591-621.

Carliner, G. (1973). "Income Elasticity of Housing Demand". *The Review of Economics and Statistics 55*(4), 528-532.

Central Bank of Hungary (2011). Quarterly Report on Inflation. Technical report, Magyar Nemzeti Bank (Central Bank of Hungary).

Cerny, A., D. K. Miles, and L. Schmidt (2010). "The Impact of Changing Demographics and Pensions on the Demand for Housing and Financial Assets". *Journal of Pension Economics and Finance 9*(03), 393-420.

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

Geanakoplos, J., M. Magill, and M. Quinzii (2004). "Demography and the Long-Run Predictability of the Stock Market". *Brookings Papers on Economic Activity 2004*(1), 241-307.

Hamermesh, D. S. (1989). "Why Do Fixed-Effects Models Perform So Poorly? The Case of Academic Salaries". NBER Working Papers 2135, National Bureau of Economic Research.

Hsiao, C. (2003). *Analysis of Panel Data* (2nd ed.). Econometric Society Monographs. Cambridge: Cambridge University Press.

Iacovou, M. and A. Skew (2010). "Houshold Structure in the EU". Methodologies and Working Papers 2010-10, Institute for Social and Economic Research, University of Essex.

Kiss, G. and G. Vadas (2007). "The Role of the Housing Market in Monetary Transmission - Evidence from Hungary". *European Journal of Housing Policy 7*(3), 299-317.

Kovács, Z. and Z. Dövényi (2006). "Budapest: The Post-Socialist Metropolitan Periphery 'Catching Up' and Individual Development Path". *European Spatial Research and Policy 13*(2), 23-41.

Li, W. and R. Yao (2007). "The Life-Cycle Effects of House Price Changes". *Journal of Money, Credit and Banking 39*(6), 1375-1409.

Lindh, T. and B. Malmberg (2008). "Demography and Housing Demand - What Can We Learn from Residential Construction Data?". *Journal of Population Economics 21*, 521-539.

Malkiel, B. G. (2010). Bubbles in Asset Prices. Working paper no. 200, CEPS.

Mankiw, N. G. and D. N. Weil (1989). "The Baby Boom, the Baby Bust, and the Housing Market". *Regional Science and Urban Economics 19*(2), 235-258.

Martin, R. F. (2006). "The Baby Boom: Predictability in House Prices and Interest Rates". 2006 Meeting Papers 84, Society for Economic Dynamics.

Meen, G. (1990). "The Removal of Mortgage Market Constraints and the Implications for Econometric Modelling of UK House Prices". *Oxford Bulletin of Economics and Statistics 52*(1), 1-23.

Meen, G. (1996). "Ten New Propositions in UK Housing Macroeconomics: An Overview of the 1980s and Early 1990s". *Urban Studies 33*(3), 425-444.

Polinsky, A. M. and D. T. Ellwood (1979). "An Empirical Reconciliation of Micro and Grouped Estimates of the Demand for Housing". *The Review of Economics and Statistics 61*(2), 199-205.

Poterba, J. M. (1984). "Tax Subsidies to Owner-Occupied Housing: An Asset-Market Approach". *The Quarterly Journal of Economics 99*(4), 729-752.

Poterba, J. M., D. N. Weil, and R. Shiller (1991). "House Price Dynamics: The Role of Tax Policy and Demography". *Brookings Papers on Economic Activity 1991*(2), 143-203.

Schaffer, M. E. (2005). XTIVREG2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models. Statistical Software Components, Boston College Department of Economics.

Song, I. H. (2010). House Prices and Consumption. MPRA Paper No. 27481, University Library of Munich, Germany.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). MIT Press Books. The MIT Press.

Appendix A: Derivation of the user cost for housing capital

In this short derivation I provide the detailed solution of the problem formulated in Meen (1990). The present-value Hamiltonian of the optimization problem defined by equations (2.1)-(2.4) is given by

$$J = e^{-\rho t}U(C(t), H(t)) + \lambda(t)[Y(t) + iA(t) - C(t) - S(t) - p_h(t)I_h(t)]+$$
$$+\mu(t)[I_h(t) - \delta H(t)] + \epsilon(t)[S(t) - \pi A(t)]. \tag{A.1}$$

The first order conditions are then obtained by differentiating J with respect to the control (C, I_h, S) and state variables (H, A):

$$\frac{\partial J}{\partial C(t)} = 0: \qquad e^{-\rho t}U'_C = \lambda(t), \tag{A.2}$$
$$\frac{\partial J}{\partial I_h(t)} = 0: \qquad \mu(t) = \lambda(t)p_h(t), \tag{A.3}$$
$$\frac{\partial J}{\partial S(t)} = 0: \qquad \lambda(t) = \epsilon(t), \tag{A.4}$$
$$\frac{\partial J}{\partial H(t)} = -\dot{\mu}: \quad e^{-\rho t}U'_H = \mu(t)\delta - \dot{\mu}(t), \tag{A.5}$$
$$\frac{\partial J}{\partial A(t)} = -\dot{\epsilon}: \quad \lambda(t)i = \pi\epsilon(t) - \dot{\epsilon}(t). \tag{A.6}$$

Differentiating (A.3) and (A.4) with respect to time yields

$$\dot{\mu}(t) = \dot{\lambda}(t)p_h(t) + \lambda(t)\dot{p}_h(t) \tag{A.7}$$
$$\dot{\lambda}(t) = \dot{\epsilon}(t) \tag{A.8}$$

Using (A.2) and (A.4) one can substitute out λ and ϵ from (A.6) and arrive at

$$e^{-\rho t}U'_C(i-\pi) = -\dot{\epsilon}(t). \tag{A.9}$$

Then by plugging (A.3), (A.6) and (A.7) into (A.5) leads to

$$e^{-\rho t}U'_H = \lambda(t)[p_h(t)\delta - \dot{p}_h(t)] - \dot{\epsilon}p_h(t). \tag{A.10}$$

Once again, using (A.2) to substitute out λ from (A.10) and expressing $-\dot{\epsilon}$ gives

$$\frac{e^{-\rho t}U'_H - e^{-\rho t}U'_C[p_h(t)\delta - \dot{p}_h(t)]}{p_h(t)} = -\dot{\epsilon}(t). \tag{A.11}$$

Finally, combining (A.11) and (A.9) gives the solution

$$\frac{U'_H}{U'_C} = p_h(t)[i + \delta - \pi - \dot{p}_h(t)/p_h(t)]. \tag{A.12}$$

Appendix B: Definition of variables

Micro-regional real house prices (P_{it}). Mean square meter price of transactions (source: National Tax and Customs Office) in the micro-region deflated with the consumer price index (source: Hungarian Central Statistical Office).

County level real house prices (P_{jt}). The following hedonic regression was estimated on transaction level data for each county:

$$\ln price_i = \alpha + (\text{controls: size}_i, \text{type}_i, \text{location}_i) + \text{year dummies} + \epsilon_i, \tag{B.1}$$

where $i = 1...N_j$ are the dwellings that were subject to a transaction between 2000 and 2010 in county j. Type controls include indicator variables for detached houses and for block of flats built from panels. Location controls include dummies for:

- the size of the city in which the dwelling is,
- being in the agglomeration of Lake Balaton,
- being in the agglomeration of Budapest.

In case of Budapest it also includes indicators for the most expensive and for the cheapest districts. The coefficients of the year dummies were then transformed to real price indices.

Housing stock (H). Stock of dwellings at the beginning of the year. Source: HCSO.

Population (POP). Continuous registration of mid-year population. Source: HCSO.

Real per capita permanent income (Y/POP). In order to capture permanent income and lagged effects I set $(Y/POP)_{it} = 0.5INC_{it}/CPI_t + 0.3INC_{i,t-1}/CPI_{t-1} +$

$0.2INC_{i,t-2}/CPI_{t-2}$, where INC_{it} is the real personal income tax basis per number of permanent inhabitants for the micro-regional data (source: Hungarian Regional Database). At the county level INC_{it} is the average real net wage multiplied by the number of full time employed and divided by population (source: HCSO).

User cost (UC). The user cost is proxied by $i - (p_{j,t-1} - p_{j,t-2})$, where i is the average mortgage rate (source: Central Bank of Hungary) and j indicates the county. Lagged house price changes are used to capture expected capital gains. The reason for using the county level changes is to reduce measurement error in UC. The average mortgage rate is the weighted mortgage rates for housing loans denominated in different currencies. The weights are the amount of initiated loans in the appropriate currencies. Note that nominal mortgage rates are used instead of real mortgage rates. The reason for this is that higher nominal rates make loans front loaded, as wages are rigid. Therefore, higher nominal rates should also have real effects.

Real construction cost ($CCost$). Gross average wages in the construction sector are adjusted with the employer's contributions. Note that this variable is only available at the county level. Source: HCSO.

Share of population aged 20 to 29 ($POP20-29$). Population as of 1st of January aged 20 to 25 divided by POP. Source: HCSO.

Appendix C: Additional estimation results

First stage results for instrumenting per capita housing stock

Table C.1: First stage results for Table 4.1

First-difference estimation Dependent variable: $(h/pop)_{i,t}$				
	IV(1)	IV(2)	IV(3)	IV(4)
$(h/pop)_{i,t-3}$	0.314	0.311	0.310	0.349
	[0.039]	[0.035]	[0.035]	[0.049]
$ccost_{i,t-3}$	-0.006	-0.006	-0.006	-0.006
	[0.004]	[0.003]	[0.003]	[0.002]
$(y/pop)_{it}$	0.022			
	[0.028]			
$p_{i,t-2}$		0.005	0.005	0.004
		[0.002]	[0.002]	[0.002]
UC_{it}	-0.003	-0.005		
	[0.003]	[0.004]		
$UC_{i,t-1}$			0.001	-0.000
			[0.003]	[0.003]
$pop20-29_{i,t-1}$				0.013
				[0.001]
Year fixed effects	Yes	Yes	Yes	Yes
Obs.	862	862	865	865
R-squared	0.176	0.176	0.174	0.325
Regions	173	173	174	174

Standard errors are clustered by micro-region. All variables are differenced and weighted by micro-regional housing stock.

Table C.2: First stage results for Table 5.1

Within estimation Dependent variable: $(h/pop)_{i,t}$	IV(1)	IV(2)	IV(3)	IV(4)
$trend_t$	0.009	0.009	0.009	0.013
	[0.002]	[0.002]	[0.002]	[0.002]
$ccost_{i,t-3}$	0.016	0.015	0.015	0.000
	[0.014]	[0.014]	[0.015]	[0.008]
$(y/pop)_{it}$	-0.013			
	[0.025]			
$p_{i,t-2}$		0.002	0.004	-0.007
		[0.005]	[0.006]	[0.008]
$UC_{i,t}$	0.017	0.016		
	[0.011]	[0.012]		
$UC_{i,t-1}$			0.004	0.002
			[0.012]	[0.013]
$pop20-29_{i,t-1}$				0.012
				[0.003]
y07	0.000	0.000	-0.001	0.000
	[0.003]	[0.003]	[0.002]	[0.002]
y08	-0.000	-0.000	-0.001	0.002
	[0.005]	[0.005]	[0.005]	[0.005]
Obs.	1037	1037	1040	1040
R-squared	0.812	0.811	0.810	0.833
Regions	174	174	174	174

Standard errors are clustered by micro-region. All variables are weighted by micro-regional housing stock.

Estimation results using different weights

Table C.3: Micro-regional results using number of transactions in 2007 as weights.

	(1)	(2)	IV(1)	IV(2)	IV(3)	IV(4)
	First-difference estimation					
	Dependent variable: micro-regional log real house price ($p_{i,t}$)					
$p_{i,t-1}$	-0.005	-0.275	-0.246	-0.594	-0.294	-0.148
	[0.056]	[0.049]	[0.058]	[0.523]	[0.174]	[0.150]
$(y/pop)_{it}$	0.795	0.530	0.385			
	[0.113]	[0.243]	[0.302]			
$(h/pop)_{it}$	-0.458	-0.090	-4.805	-4.846	-3.737	-1.943
	[0.518]	[0.512]	[1.057]	[2.515]	[1.072]	[0.779]
$pop20-29_{it}$						0.069
						[0.031]
UC_{it}	-0.253	-0.191	-0.176	-0.349	-0.124	-0.105
	[0.033]	[0.047]	[0.048]	[0.254]	[0.141]	[0.139]
Year fixed effects	No	Yes	Yes	Yes	Yes	Yes
Obs.	1034	1034	861	861	861	861
R-squared	0.169	0.450	0.012	0.215	0.317	0.351
Regions	173	173	173	173	173	173

Note: standard errors in brackets are based on the clustered covariance matrix by micro-regions. All variables are differenced (including year fixed effects and instruments) and weighted by the square root of transactions. y/pop - log per capita permanent income, h/pop - log per capita stock of dwellings, UC - user cost, $pop20-29$ - percent of population aged between 20 and 29. In IV(1) $(h/pop)_{it}$ assumed to be endogenous (instruments: $(h/pop)_{i,t-3}$, $ccost_{i,t-3}$), in IV(2) h/pop_{it} and $p_{i,t-1}$ assumed to be endogenous (instruments: $(h/pop)_{i,t-3}$, $ccost_{i,t-3}$, $p_{i,t-2}$), in IV(3) $(h/pop)_{it}$, UC_{it} and $p_{i,t-1}$ assumed endogenous (instruments: $(h/pop)_{i,t-3}$, $ccost_{i,t-3}$, $p_{i,t-2}$, UC_{t-1}) and in IV(4) $(h/pop)_{it}$, UC_{it}, $p_{i,t-1}$ and $pop20-29_{i,t}$ assumed endogenous (instruments: $(h/pop)_{i,t-3}$, $ccost_{i,t-3}$, $p_{i,t-2}$, UC_{t-1} and $pop20-29_{i,t-1}$).

Table C.4: County level results using number of transactions in 2007 as weights.

First-difference estimation						
Dependent variable: county level log real house price ($p_{j,t}$)						
	(1)	IV(1')	IV(2')	IV(3')	IV(4')	IV(5)
$p_{j,t-1}$	0.394	0.369	0.233	0.311	0.040	0.036
	[0.066]	[0.069]	[0.117]	[0.116]	[0.067]	[0.066]
$(y/pop)_{jt}$	0.669	0.500	0.703	0.415		
	[0.134]	[0.156]	[0.191]	[0.212]		
$(h/pop)_{jt}$	-2.436	-2.979	-3.394	-2.637	-2.719	-2.388
	[0.246]	[0.346]	[0.289]	[0.374]	[0.524]	[0.839]
$pop20-29_{jt}$					0.034	0.044
					[0.011]	[0.019]
$crisis_t$				-0.039	-0.007	-0.005
				[0.009]	[0.008]	[0.007]
$mrate_t$	-1.108	-0.867	-0.583	-0.419	-0.160	-0.202
	[0.233]	[0.266]	[0.200]	[0.190]	[0.162]	[0.176]
Obs.	160	140	140	140	140	140
R-squared	0.728	0.714	0.701	0.736	0.544	0.543
Counties	20	20	20	20	20	20

Note: standard errors in brackets are based on the clustered covariance matrix by counties. All variables are differenced (including year fixed effects and instruments) and weighted by the square root of transactions. y/pop - log per capita permanent income, h/pop - log per capita stock of dwellings, $mrate$ - mortgage rate, $pop20-29$ - percent of population aged between 20 and 29, $crisis$ - 1 if $t = 2009$ or $t = 2010$, 0 otherwise. In IV(1') $(h/pop)_{jt}$ assumed to be endogenous (instruments: $(h/pop)_{j,t-3}$, $ccost_{j,t-3}$), in IV(2')-IV(4') h/pop_{jt} and $p_{j,t-1}$ assumed to be endogenous (instruments: $(h/pop)_{j,t-3}$, $ccost_{j,t-3}$, $p_{j,t-2}$) and in IV(5) $(h/pop)_{it}$, $pop20-29_{jt}$ and $p_{i,t-1}$ assumed endogenous (instruments: $(h/pop)_{j,t-3}$, $ccost_{j,t-3}$, $p_{j,t-2}$, $pop20-29_{j,t-1}$).

Table C.5: Micro-regional results using population in 2000 as weights.

First-difference estimation						
Dependent variable: micro-regional log real house price ($p_{i,t}$)						
	(1)	(2)	IV(1)	IV(2)	IV(3)	IV(4)
$p_{i,t-1}$	-0.116	-0.354	-0.329	-0.293	-0.285	-0.183
	[0.050]	[0.045]	[0.055]	[0.225]	[0.149]	[0.135]
$(y/pop)_{it}$	0.894	0.528	0.416			
	[0.124]	[0.255]	[0.325]			
$(h/pop)_{it}$	-0.459	0.156	-4.421	-3.310	-3.282	-1.780
	[0.461]	[0.501]	[1.144]	[1.188]	[1.027]	[0.810]
$pop20-29_{it}$						0.082
						[0.033]
UC_{it}	-0.275	-0.218	-0.199	-0.192	-0.182	-0.137
	[0.037]	[0.047]	[0.051]	[0.111]	[0.183]	[0.189]
Year fixed effects	No	Yes	Yes	Yes	Yes	Yes
Obs.	1034	1034	861	861	861	861
R-squared	0.148	0.392	0.054	0.278	0.278	0.290
Regions	173	173	173	173	173	173

Note: standard errors in brackets are based on the clustered covariance matrix by micro-regions. All variables are differenced (including year fixed effects and instruments) and weighted by the square root of population. y/pop - log per capita permanent income, h/pop - log per capita stock of dwellings, UC - user cost, $pop20-29$ - percent of population aged between 20 and 29. In IV(1) $(h/pop)_{it}$ assumed to be endogenous (instruments: $(h/pop)_{i,t-3}$, $ccost_{i,t-3}$), in IV(2) h/pop_{it} and $p_{i,t-1}$ assumed to be endogenous (instruments: $(h/pop)_{i,t-3}$, $ccost_{i,t-3}$, $p_{i,t-2}$), in IV(3) $(h/pop)_{it}$, UC_{it} and $p_{i,t-1}$ assumed endogenous (instruments: $(h/pop)_{i,t-3}$, $ccost_{i,t-3}$, $p_{i,t-2}$, UC_{t-1}) and in IV(4) $(h/pop)_{it}$, UC_{it}, $p_{i,t-1}$ and $pop20-29_{i,t}$ assumed endogenous (instruments: $(h/pop)_{i,t-3}$, $ccost_{i,t-3}$, $p_{i,t-2}$, UC_{t-1} and $pop20-29_{i,t-1}$).

Table C.6: County level results using population in 2000 as weights.

First-difference estimation						
Dependent variable: county level log real house price ($p_{j,t}$)						
	(1)	IV(1')	IV(2')	IV(3')	IV(4')	IV(5)
$p_{j,t-1}$	0.391	0.371	0.255	0.326	0.026	0.022
	[0.061]	[0.063]	[0.123]	[0.122]	[0.066]	[0.067]
$(y/pop)_{jt}$	0.640	0.425	0.595	0.334		
	[0.131]	[0.137]	[0.190]	[0.214]		
$(h/pop)_{jt}$	-2.299	-3.012	-3.299	-2.624	-2.835	-2.565
	[0.218]	[0.303]	[0.276]	[0.374]	[0.465]	[0.738]
$pop20-29_{jt}$					0.032	0.041
					[0.011]	[0.019]
$crisis_t$				-0.037	-0.001	0.000
				[0.011]	[0.009]	[0.008]
$mrate_t$	-1.188	-0.935	-0.681	-0.543	-0.173	-0.203
	[0.248]	[0.259]	[0.206]	[0.207]	[0.177]	[0.194]
Obs.	160	140	140	140	140	140
R-squared	0.715	0.686	0.673	0.707	0.473	0.475
Counties	20	20	20	20	20	20

Note: standard errors in brackets are based on the clustered covariance matrix by counties. All variables are differenced (including year fixed effects and instruments) and weighted by the square root of population. y/pop - log per capita permanent income, h/pop - log per capita stock of dwellings, $mrate$ - mortgage rate, $pop20-29$ - percent of population aged between 20 and 29, $crisis$ - 1 if $t = 2009$ or $t = 2010$, 0 otherwise. In IV(1') $(h/pop)_{jt}$ assumed to be endogenous (instruments: $(h/pop)_{j,t-3}$, $ccost_{j,t-3}$), in IV(2')-IV(4') h/pop_{jt} and $p_{j,t-1}$ assumed to be endogenous (instruments: $(h/pop)_{j,t-3}$, $ccost_{j,t-3}$, $p_{j,t-2}$) and in IV(5) $(h/pop)_{it}$, $pop20-29_{jt}$ and $p_{i,t-1}$ assumed endogenous (instruments: $(h/pop)_{j,t-3}$, $ccost_{j,t-3}$, $p_{j,t-2}$, $pop20-29_{j,t-1}$).