Demographic Cycle, Migration and Housing Investment: a Causal Examination

By Eric Monnet and Clara Wolf*

Abstract

This paper argues that residential investment is much more volatile than GDP because the demographic variable that determines specifically housing demand (growth of population aged 20-49) is itself highly volatile. First, we depart from the literature on housing cycles which has mainly focused on explaining the cycle of housing prices by changes in financial conditions. Instead, we study the cycle of real residential investment as a share of real GDP in 20 OECD countries since 1980, and show that it is closely associated with the growth dynamics of population aged 20-49. Second, the paper develops a new method to uncover the causal effect of the growth of population aged 20-49 on housing construction. We use past demographic data as an instrument to avoid the potential endogeneity bias between migrations and the housing cycle. The instrument is strong in countries where net migration is low. Overall, we find that a 1% increase in the population aged 20-49 increases the ratio of residential investment to GDP by 1.3 pp. Demographic changes appear to be a better predictor of the cycle of the residential investment rate than any other macroeconomic variable we control for. (*JEL* E32, J11, R21)

Residential investment is one the most volatile components of GDP and usually leads or is coincident with total output (Davis and Heathcote (2005), Leamer (2007), Kydland, Rupert and Sustek (2012)). It is strikingly five times more volatile than consumption in OECD countries, although it is also mainly produced by households. The standard explanation of such a peculiar volatility is that housing is an asset. Thus, theoretical and empirical studies mostly look at the cycle of prices of existing dwellings and how it is determined by financing conditions (Bénétrix, Eichengreen and O'Rourke (2012), Jordà, Schularick and Taylor (2015), Favara and Imbs (2015), Justiniano, Primiceri and Tambalotti

*Monnet: Banque de France, DGEI-DCPM-DIACONJ, 31, rue Croix des Petits Champs, 75049 Paris, France (e-mail: eric.monnet@banque-france.fr); Wolf: Sciences Po, 28 Rue des Saints-Peres, 75007 Paris, France (email: clara.wolf@sciencespo.fr). We are grateful to Yannick Kalantzis and Nicolas Coeurdacier for their insightful comments. The views expressed herein are those of the authors and do not necessarily reflect those of the Banque de France.

(2015)). From an accounting point of view however, sales and purchases of existing dwellings do not contribute to residential investment and GDP. And there are many reasons to believe that changes in housing prices does not perfectly reflect the flows of new housing, either because of market imperfections, measurement errors, or because housing is a durable good and prices respond asymmetrically to shocks (Glaeser and Gyourko (2005)). Then, what drives real investment in housing? Beyond credit and asset prices, what about real factors? The most intuitive determinant of the number of new dwellings is demography: the variation of population (especially people who are old enough to invest but young enough to need a new house) affects directly the demand for housing. Following Mankiw and Weil (1989), macroeconomic debates on this issue have been centered on few individual country cases, and on the effect of the age structure, or of the size of age groups, on housing prices. Very few cross-country studies have investigated whether demographic variables do explain the housing construction cycle of advanced economies and the common wisdom remains that - at a national level - population growth may affect the long run housing stock but is not a main determinant of housing cycles. Furthermore, the few cross-country studies that looked for an effect of the relative size of age groups, either on housing construction (Lindh and Malmberg (2008)) or on housing prices (Agnello and Schuknecht (2011), Takáts (2012), Bracke (2013)), have not addressed the potential endogeneity of population growth to housing conditions. The growth dynamics of the population is potentially endogenous to housing investment and prices, especially if housing is indeed the business cycle (Leamer (2007)): people would tend to immigrate to countries with a more dynamic housing sector and higher GDP growth, and residents would emigrate when a country suffers from low growth (which can be due or correlated to an housing crisis).

In this paper, we make two main contributions. First, based on a panel of 20 OECD countries since 1980, we show that the growth of population aged 20-49 (which is the age range most likely to affect housing demand) is actually highly cyclical and is strongly correlated to residential investment (as a share of GDP). Such a relationship is not observed for the other components of GDP and is very robust when controlling for other parameters such as revenue per capita, unemployment, financial factors (credit and interest rates) and real house prices. Building on the seminal work of Mankiw and Weil (1989), we interpret this effect as a pure demand effect since the age profile is not expected to influence housing supply in the same way. Moreover, looking at the ratio of residential investment to GDP is a way to isolate how the growth of an age group affects housing investment differently from the other components of GDP.

Second, we propose a simple method to isolate and estimate the exogenous effect of the growth of population aged 20-49 on real residential investment.[1] We use the growth rate of the population aged 0-29, twenty years before, to predict the current variation of the population aged 20-49. The rationale of this instrument is that demographic data determined 20 years before is unlikely to be affected by the current housing cycle. According to reduced form and IV estimates, the effect is positive and significant.

By nature however, our instrument would be weak in the presence of high migration. The instrument is strong only when there is a relatively small difference between the actual growth rate of the population aged 20-49 in year t and the growth rate of the population aged 0-29 in year $t-20$, such that the latter is a good predictor of the former because the net migration in the previous 20 years was low. To cope with this issue and to discuss in what extent our previous result would be affected by a weak instrument problem, we split our sample of 20 countries in two equal sub-samples of 10 countries (in order to ensure comparability of degrees of freedom between the subsamples) which differ by the average magnitude of net migration flows.[2] According to the standard robust test for weak instruments in the presence of serial correlation and heteroskedasticity (Olea and Pflueger (2013)), the instrument is indeed strong in the sample with low migration, but weak in the sub-sample with high migration. Accordingly, we find a causal significant effect of the growth rate of population aged 20-49 only in the subsample with low migration, whereas a significant correlation is observed in both subsamples. The difference between the two groups highlights that there is a potential sizeable endogeneity bias in the presence of high migration.

In the subsample where we are able to identify a causal effect with a strong instrument, we find that when the growth rate of the population aged 0-29 in year $t-20$ increased by 1pp, the ratio of residential investment to GDP

[1] This method also circumvents the limitation of data on migrations by age which are notoriously imprecise and are usually not available.

[2] In the remaining of the paper, we always mean "net migrations" when we refer to high or low migration flows. Talking about these net migration flows, we do not distinguish between flows of citizens and foreigners. For example, a Italian citizen who left Italy at 20, became a Japanese resident but came back to Italy at 35, would count as an Italian emigrate at 20 and as an immigrate to Italy at 35.

in year t increases by 0.92pp (reduced form specification).[3] The instrumental variable estimator leads to a slightly stronger effect: when the growth rate of the population aged 20-49 in year t increases by 1pp, the ratio of residential investment to GDP in year t increases by around 1.3 pp.

It is important to note that although we identify a general causal effect of population growth on housing, this paper cannot distinguish between exogenous and endogenous effects of migrations and estimate their effects. Despite such a limitation, our method has direct and important implications for macroeconomic policies and modeling, as well as for forecasts and debates on future GDP growth and the housing market.

In more intuitive terms, the conclusion of our study is that a significant part of the housing construction cycle of the 2000s in OECD countries (both the boom of the early 2000s and the bust of the late 2000s, which were amplified by financial factors) would have been predictable using past demographic data, except in countries where migration was so high that it made such predictions impossible. Thus, our paper also contributes to the large literature on the housing boom that preceded the 2007 financial crisis in many countries (see Caldera and Johansson (2013) and Bénétrix, Eichengreen and O'Rourke (2012) for reviews).[4] The overtime pattern of demographic data and housing cycles have been very similar across countries: within-country variations overtime are more important than cross-country variations for a given year.[5] Overall, the (mostly predictable) slowdown or decrease of the growth of the population aged 20-49 in OECD countries in the second half of the 2000s stands as an explanation of the slowdown of residential investment and GDP growth during the Great Recession. Furthermore, the current age structure is such that the United Nations predict a 4% decrease on average, of the numbers of people aged 20 - 49 between 2015 and 2045 in the 20 OECD countries included in our study.[6] Our findings suggest that these demographic changes are likely to have a strong

[3] The coefficient of the reduced form equation should be interpreted as the impact of the predicted growth of population aged 20-49, without migrations (and with a constant mortality rate by age).

[4] Data on residential investment and our simple decomposition of demographic data by country also shed a new light on the different cycles of housing construction of the last 20 years in the OECD. In Ireland and Spain, the notorious housing boom of the early 2000s was correlated to a strong increase in the population aged 20-49, which was not predictable using past demographic data but due to immigration.

[5] See Figure B1 and Figure B3 in Appendix B.

[6] See Figure B2 in Appendix B.

deflationary impact on residential investment, although it can potentially be mitigated by migrations.[7]

Because of our focus on demography, our paper is also related to a small literature that has shown that, when accounting for specific factors of the housing market, easy credit conditions and low interest rates are not the dominant factor of housing booms. Glaeser, Gottlieb and Gyourko (2012) raised such a conclusion on the US as they observe a low correlation between interest rates and building permits at the national level.[8] Our macroeconomic cross-country investigation leads to the conclusion that when demographic and income factors are appropriately taken into account, the role of finance is considerably reduced. In our estimations, the growth of real credit is not significant. The effect of real interest rates is significant (but very small) only when the endogeneity of demographic variables is not appropriately taken into account ; it is insignificant in the reduced form and IV estimations. Demographic changes appear to be a better predictor of the cycle of the residential investment rate than any other macroeconomic variable we control for.

Finally, understanding better the determinants of the pattern of residential investment is very important for macroeconomic modeling. As highlighted by Davis and Heathcote (2005), Iacoviello and Neri (2010), business cycle models with technology shocks are unable to reproduce the dynamic of housing investment over the business cycle. The only way to reproduce the dynamic of housing investment is to simulate shocks on housing demand (Iacoviello and Neri (2010)).[9] The direct implication of our paper for the DSGE literature is that macroeconomic models should incorporate shocks on the demographic structure of the population (or interpret house preference shocks as such demographic shocks) in order to account for the peculiar dynamic of residential investment over GDP. Our finding that the growth of population aged 20-49 influences the ratio of residential investment over GDP, but not consumption or business investment over GDP, clearly highlights the peculiarity of the age profile of the demand of housing. Compared to other age groups, people aged

[7] Assuming that long-term supply housing constraints would remain stable and that the demand for housing (which depend on demographic factors highlighted in this paper) would still be met by a supply.

[8] Shiller (2007) also shows weak historical evidence on the correlation between interest rates and housing prices.

[9] In DSGE models, housing investment is either modeled as home production (Benhabib, Rogerson and Wright (1991)) or as an equilibrium output determined by consumption of housing by households and production of housing by firms (Iacoviello and Neri (2010)). Common to these different macroeconomic approaches is the emphasis on the similarity between consumption and housing investment.

20-49 spend more in residential investment that they do in consumption or that they save to invest in production.

The paper is structured as follows. Section 1 explores the several caveats of the past literature and highlights our contribution. Section 2 presents the benchmark estimations and main results. Section 3 tackles the issue of endogeneity of the 20-49 age group's growth to housing conditions and explains our IV strategy. Section 4 provides a more thorough analysis of migrations, by age, overtime and by countries, and their correlations with the cycle of construction, to assess the potential size of the endogeneity bias.

I. Related Literature

A. *Quantities and prices*

Residential investment is one of the most volatile components of GDP. It is 4.4 times more volatile than GDP whereas non-residential investment is only 3.5 times more volatile and consumption is less volatile than GDP (0.9 times GDP's volatility).[10] The standard explanation of such a peculiar pattern is that housing is an asset (Topel and Rosen (1988)). As a consequence, studying the behavior of asset (house) prices is supposed to provide an explanation of the volatility of housing flows. Thus, most of the studies on the housing cycle look at the determinants of housing prices. Recent work on this topic - motivated by the US subprime mortgage crisis and the global financial turmoils - has focused mainly on how financial factors shape the cycle of housing prices (Bénétrix, Eichengreen and O'Rourke (2012), Jordà, Schularick and Taylor (2015), Favara and Imbs (2015), Justiniano, Primiceri and Tambalotti (2015)).

An older strand of literature, starting with Mankiw and Weil (1989), focused on explaining prices by demographic changes. The argument is that, in a long term perspective, housing supply should be very elastic. Hence, persistent changes in demand, such as demographic shifts, will be reflected in the price or quantity of housing. This literature is rather inconclusive and results depend on the countries and time span studied. Some papers find a significant link between demographics and housing prices (Mankiw and Weil (1989) Ohtake and Shintani (1996), Takáts (2012)) whereas some authors find that demographics is of minor importance or has no impact (Engelhardt and Poterba (1991), Peek and Wilcox (1991), Holland (1991)). From a business cycle analysis perspective,

[10] Only the stocks component and net exports are more volatile. These figures (in real terms) are for our sample of 20 OECD countries over the period 1980-2014. See the description of data in the Appendix A.

this literature suffers from several caveats. Either it is centered on housing prices without justifying this focus (Peek and Wilcox (1991), Takáts (2012), Eichholtz and Lindenthal (2014), Green and Hendershott (1996)) or it looks at housing investment and prices indifferently (Mankiw and Weil (1989) , Lee et al. (2001), Lindh and Malmberg (2008)) without explaining the differences between these variables. However, we believe that there are several reasons to focus on residential investment rather than on housing prices, if we want to explain the cycle of the new flows that improve the housing stock and contribute to GDP.

First, cross-country empirical studies systematically use prices of existing dwellings. The main reason is that prices of existing dwellings are more easily registered and thus more widely available than prices of new dwellings (see Mack and Martínez-García (2011) and Knoll, Schularick and Steger (2014) for recent surveys on the sources for housing prices). Only the US and Canada have data on prices of new dwelling for a long time span. Most countries, especially European ones, have started to publish official, nationwide, series of prices of new dwellings, much later, usually after 2005. Other, as Australia, publish an index for existing and new dwelling without making the distinction possible. Since existing and new dwellings have different supplies and may face different demand, there is no reason for the cycle of prices to be the same in these two sectors, even if both prices are likely to converge in the medium and long run.

Second, available information on house prices concerns the private sector only. For a number of countries, it is a severe limitation, as a significant part of the population lives in the social housing sector. According to OECD data (Caldera and Johansson (2013)), social housing accounts for the majority of the rental sector in the Netherlands, Austria, the Nordic countries, the United Kingdom and Ireland. In these countries, as well as in France, the social housing sector's size is more than 15% of the total dwelling stock. French data - for which a detailed decomposition of housing investment is available - shows that 10 to 20% of residential investment is due to social housing and that this share varies a lot over the business cycle: it increases when the private sector investment decreases (Faubert, Monnet and Sutter (2015)).

Third, cycles of quantities and prices can differ in a significant way. In the US case, Leamer (2007, 2015) argues that "homes experience a volume cycle, not a price cycle". Indeed, in a long-term perspective, the cycle of prices is milder than the cycle of residential investment. In our sample, residential investment is 1.6 times more volatile than real housing prices (see Figure B6). It might be due to the fact that, as described previously, we only observe the price of existing dwellings for half of our sample whereas we look at the quantities of new constructions, and that the demand for new dwellings varies more than the one for existing dwellings (if construction costs are low for example). It also might be due, as argued by Leamer, to “pathologies in the price discovery process”: prices are sticky because of incomplete information (say, for example, that the seller thinks that the drop in demand is only temporary and thus keeps the price stable whereas the drop is permanent), which amplifies the volume cycle, as high prices deter sales. Another argument - which does not rely on imperfect information or substitution - is given by Glaeser and Gyourko (2005) in their study of the dynamic of cities. Taking into account the peculiar characteristics of the housing market (i.e housing is a durable good whose quantity increases with higher demand but cannot be reduced by lower demand), the authors emphasize the profound asymmetric response of prices to positive and negative demand shocks. Without legal or geographic constraints on supply, housing prices will increase less than residential investment during housing demand booms (because new supply is elastic when prices are at or above construction costs), but will decrease in a similar way during demand busts.

As a conclusion, there are several arguments to claim that observed housing price can be a very imperfect indicator of the dynamic of residential investment. Furthermore, because of the imperfectness of prices data, residential investment is a more consistent and comparable statistics when comparing housing cycles across countries.

B. *Supply and demand*

Mankiw and Weil (1989) use the US census to show that the purchase of an house is age specific, from which they conclude that the age profile strongly affects housing demand.[11] As we rely on their seminal contribution, it is worth explaining further why demographic variables such as the population’s age distribution allow us to identify demand from supply of housing. The identification is based on the fact that demand for housing is determined by the growth of a specific age group only (people aged 20-49 in this paper, cf. Section II) whereas supply of new housing does not depend on the age structure of the population. In other words, population aged 20-49 is more likely to demand new housing investment than the population aged 50 +, but not likely to supply more land for housing construction. It also should be noted that the

[11] The age profile of house purchases is well documented in other countries too. For example, Faubert, Monnet and Sutter (2015) display data from the French 2004 census showing that 75% of recent buyers are aged between 20 and 49. Note, however that in the censuses, it is impossible to distinguish between buyers of existing dwellings and buyers of new dwellings.

identification is reinforced by the fact that suppliers of new housing differ from the suppliers of existing dwellings. Suppliers of existing dwellings may include a significant proportion of middle-aged people selling a home to buy a new one. The demanders and the suppliers of existing dwellings may share similar properties (age, income etc.). On the contrary, suppliers of new housing are not the same individuals than the ones who look for new housing. First, the supply of new dwellings is made up of both the sellers of land and the real estate companies that build the new houses. They may face different incentives and (legal or budget) constraints. Second, the sellers of land are private individuals but also private firms as well as the State or other public institutions. Thus, there is no evidence that the supply of land would be determined by the age structure of the population in the same way as demand is. Hence, using the age profile to identify a pure housing demand effect performs better in the case of housing investment than in the case of transactions or prices of existing homes.

Previous literature considers population as a long-run determinant of housing demand (with the noticeable exception of Muellbauer and Murphy (1997)) whereas we find that an age group growth can substantially vary and comove with residential investment over time (as the population aged between 20 and 49 years old, see Figure 1 for some countries, and Figure B4 for the whole sample; and especially for the 30-49 years old, see Figure B5). Mankiw and Weil (1989) built an index of the level (rather than the change) of housing demand weighting the number of households by their probability to buy (depending on age). Lindh and Malmberg (2008) divide the number of people of each age group by the total population.[12] Takáts (2012) uses the dependency ratio (the number of young people divided by the number of old people). Surprisingly, the growth rate of age groups has not been used as an explanatory variable in previous papers on housing and demography. Hence, these previous studies have been unable to capture the high volatility of housing demand. The high cyclicality of the 20-49 age group's growth is a key fact highlighted in this paper, in order to explain that a pure demand effect - based on fundamentals - might provide a sufficient explanation of the high volatility of residential investment.

Finally, as we will see in the Section III, none of the previous papers on housing and demography has raised the potential issue of endogeneity between residential investment and population flows.

[12] Nevertheless, this specification induces high multicollinearity between the coefficients of age groups' shares, and an age group must be dropped because the sum of coefficients is equal to one.

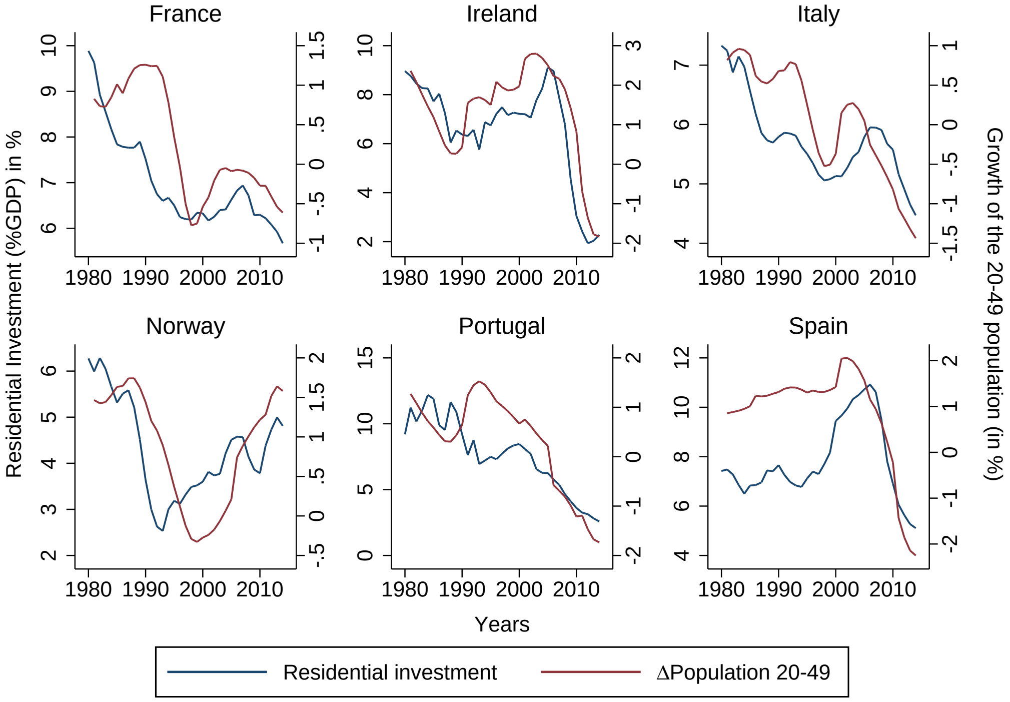


FIGURE 1. RESIDENTIAL INVESTMENT AND THE 20-49 POPULATION GROWTH
FOR THE WHOLE SAMPLE, SEE FIGURE B4

II. Housing Investment and Demographic Cycles

In this section, we present the data used in our study, we explain our specification choices, and study the correlation between residential investment and different growth rates of age groups.

A. *Data*

In National Accounts, residential investment, also named Gross fixed capital formation (GFCF) in dwellings, excludes other constructions (such as non residential buildings or infrastructures). It includes GFCF in construction of households, GFCF in social housing (usually registered as corporate investment in national accounts), and the few GFCF in dwellings of the public sector (some public residence for students, for example). GFCF in dwellings include both construction of new homes and renovation/major maintenance work of existing homes. Unfortunately, the published statistics do not distinguish between these two components.[13] Expenditures in improvements and renovations are registered as investment only if they increase the value of the building. Otherwise, they are registered as consumption (in construction).

[13] It also includes property-related fees and services, which are a very small part of the total. These fees and services (architects, notaries) are mainly associated with new construction, but also with existing dwellings when applicable (i.e only for notaries fees).

How does GFCF in dwellings differ from other measures of housing construction? Building permits and housing starts are sometimes used to study the cycle of housing construction, especially in the US where a long and consistent time series of housing starts is available (Leamer (2007), Glaeser, Gottlieb and Gyourko (2012)). However, these series can differ substantially from real housing investment for three main reasons. First, these are usually the number of new permits and housing starts (although some countries publish the figures in square meters), without any information on the value of the expenditures associated with each new dwelling. The average amount of investment per new home can vary greatly overtime. Second, building permits can be cancelled, and the cancellation rate typically varies over the business cycle (for example, in France it was 10% on average between 2000 and 2007 during the housing construction boom, but 25% from 2012 to 2015 during the bust). Third, building permits and housing starts include only a small part of the actual improvement of existing homes, and the regulation on the minimum size of construction requiring a building permit may vary both overtime and across countries.

As a conclusion, the best way to measure the quantity of investment that really improves the value of the (both private and public) housing stock overtime is to focus on the GFCF in dwellings. From a more practical point of view, long term GFCF series are much more easily available overtime with a consistent definition within and between countries. For most countries, series of permits or housing starts are much shorter or suffer from important breaks (often due to changes in the regulation of building permits).[14]

In our study, we scale the residential investment by GDP. To compute this ratio, we use gross fixed capital formation (GFCF) in dwellings in 2010 constant prices, divided by the gross domestic product (GDP) at market prices in 2010 constant prices, both taken from the AMECO database. Sources of other macroeconomic variables used in the next sections are described in depth in the Appendix A . They can be briefly summed up as follows. The population by age group is taken from the World Population Prospect of the United Nations and represent the population as of the 1st of July for each year. Both the real personal disposable income and the real housing prices were taken from the consolidated database of the Dallas Fed. The real long-term interest rate is the long-term (10 years) interest rate on government bonds taken from OECD Economic Outlook deflated by the CPI (Consumer Price Index) taken from OECD also. Credit to the private non-financial sector from all sectors, as

[14] GFCF in construction of households which provides a measure of the investment in private housing stock is also less available than the GFCF in dwellings.

well as credit to households, are taken from the BIS database and expressed in market value, in domestic currency. We deflate them using the CPI from the OECD to obtain real credits. Finally, the unemployment rate as percentage of civilian labour force is taken from the AMECO database. Descriptive statistics of these variables are presented in Table 1.

Table 1 – Summary Statistics

	Mean	Standard error			N
		Overall	Between	Within	
Residential Investment (%GDP)	6.16	2.17	1.57	1.53	700
Population growth (1980-2014)					
0-19	−0.37	0.99	0.62	0.79	680
20-49	0.55	0.90	0.47	0.77	680
50 plus	1.45	0.76	0.49	0.59	680
Population growth (1960-1994)					
0-29	0.18	0.93	0.42	0.83	680
30 plus	1.11	0.63	0.50	0.40	680
Migration balance of the 20-49	0.52	0.65	0.43	0.50	520
Natural balance of the 20-49	−0.09	0.77	0.38	0.67	520
Controls					
Real PDI growth	1.69	2.95	1.16	2.72	680
Real Housing Price growth	1.46	6.54	1.35	6.41	680
Real Long-term Interest Rate	1.41	5.27	0.62	5.24	700
ΔUnemployment	0.12	1.08	0.16	1.07	680
Real Credit Growth	4.26	4.93	1.31	4.76	646
Real Household Credit Growth	4.97	5.33	1.45	5.14	429

Note: Mean value of country-year pooled data of the sample of 20 OECD countries between 1980 and 2014. All variables are expressed in percentage or percentage points (for the interest rate and the unemployment). Migration and natural balance are expressed in percentage of the population aged between 20 and 49 years old.

B. *Specification*

In order to study the correlation between housing investment and changes in age groups, we estimate a balanced panel on 20 countries with annual data from 1980 to 2014. We take a reasonably homogenous sample of OECD countries[15] to minimize problems of model heterogeneity. Nevertheless, there are still large disparities in term of reglementation on housing and mortgage markets between our countries. To account for this heterogeneity, we include

[15] The sample includes Austria, Australia, Belgium, Canada, Switzerland, Denmark, Spain, Finland, France, UK, Greece, Ireland, Italy, Japan, Luxembourg, Netherlands, Norway, Portugal, Sweden and US. We exclude Germany since there is a problem with data homogeneity before and after the reunification (residential investment is only available for West Germany before 1991).

country-fixed effects. We clustered our standard errors by country to correct for serial correlation and heteroskedasticity in the residuals.[16] To avoid the risk of spurious regression , we check the stationnarity of our variables.[17]

To account for the impact of the growth of age groups on residential investment, we estimate :

$$GFCF\ dwellings/GDP_{c,t} \quad = \quad \alpha_0 + \sum_{i=1}^{3} \alpha_i \Delta Pop\ age\ group_{c,t,i} + \alpha_4 X_{c,t} + \gamma_c + \xi_{c,t}$$

where $GFCF\ dwellings/GDP_{c,t}$ is the residential investment as a share of GDP (in percentage) in country c at time t, $\Delta Pop\ age\ group_{c,t,i}$ are the growth (in percentage) of the 3 age groups, respectively aged 0-19, 20-49 and 50 years old and above. $X_{c,t}$ includes several time-varying controls : the growth (in percentage) of the real personal disposable income per capita, real housing prices, and real credit, the variation of unemployment and the level of real long-term interest rates. The fixed effects (γ_c) ensure that all the country-specific influences are accounted for, provided they are invariant over time, and $\xi_{c,t}$ is a disturbance term.

1. Dependent variable

There are two reasons to use the ratio of residential investment over GDP as a dependent variable rather than the level of residential investment. First, as all the main expenditure components of GDP, residential investment has a unit root. There is a high risk of spurious correlations when using such a variable as a dependent variable.

The second reason is due to the focus of this paper on demographic factors and, more especially, on the impact of the growth of population aged 20-49 on housing investment. Since population growth is closely linked to GDP growth, both for theoretical and accounting reasons, there is a mechanical positive impact of the growth of all age groups on any component of the GDP. Then, dividing residential investment by GDP is a way to isolate the specific economic impact of the growth of age groups on residential investment from the mechanical effect of population growth on all components of GDP. A positive impact of the growth of an age group on the ratio of residential investment to GDP should be interpreted as a specific positive effect on residential investment

[16] The clustering is needed, since our residuals present evidence of serial correlation and heteroskedasticity, see the results of the tests in Table C1 in Appendix C.

[17] For the results of the tests, see Table C3 in Appendix C.

taking into account the positive effect of this age group on the other components of GDP. If the growth of an age group affects positively all the components of GDP (including residential investment) by a similar magnitude, then the ratio of residential investment will nonetheless decrease. If the growth of an age group affects positively residential investment only, then the ratio of residential investment will increase, all other things being equal.

As already shown by Kydland, Rupert and Sustek (2012) for 6 countries, we find that for the 20 countries in our sample, the evolution of residential investment is almost coincident to GDP. The volatility of residential investment is 4.4 times GDP's volatility. It implies that the value of the residential investment rate is mainly driven by the fluctuations of residential investment. As shown in Table 4, the periods of expansions of the residential investment to GDP ratio are associated (in average) with positive growth rates of residential investment whereas the periods of a decrease of this ratio are associated with negative growth of residential investment. GDP growth, on the contrary, is positive during periods of both expansion and recession of residential investment to GDP, although slightly lower during housing recessions. As seen on Table 4 and discussed below (Section C), consumption and non-residential investment follow different patterns.

2. Age groups

The choice of the number of age groups entails a trade-off. With 5-year age groups, the age group's coefficients won't be precisely estimated since age groups are rather correlated. In the literature, the common solution is to use one variable aggregating all the information of the distribution such as dependency ratios (Takáts (2012)) or by imposing a polynomial restriction on the coefficients such as Fair and Dominguez (1991) or Higgins (1998), or in the case of Mankiw and Weil (1989) by imposing coefficients that have been estimated using census cross-sectional data. However, Mankiw and Weil report slightly different age profiles for 1970 and 1980 census, invalidating a posteriori the restriction imposed. Such restrictions discard some information on relative movements in the distribution that could be important.

To choose the number of age groups, we estimate the relation for 6 age groups and aggregate the age groups that have a similar impact on residential investment (cf. Table 2), and check whether multicollinearity is an issue. Based on these estimation results, we divide the population in three age groups : the 0-19 years old (the children, who do not take economic decisions regarding housing), the 20-49 years old (the adult group in which household formation takes place), and the 50 years old and above (middle-aged people and retirees).

3. Controls

Theoretically, the demand of housing should increase with income growth. The personal disposable income represents the income resources "at hand" for the households for current purchases, expressed per working-age population. It is a common measure to assess the affordability of housing (Mack and Martínez-García (2011)).[18] Since residential investment is expressed as a share of GDP, an increase in the growth of the real personal disposable income per capita leading to an increase in the residential rate, would mean that households spend more on housing than on other goods. In other words, new dwellings or renovation would be superior goods.

To control for the credit market conditions, we use the real long-term interest rates, i.e. the nominal rates on 10-year sovereign bonds deflated by the consumer price indexes, which are the reference rate for housing mortgages, as well as the real growth rate of credit to the private non financial sector, as in Agnello and Schuknecht (2011) and Bracke (2013). Real credit to the private non financial sector includes credit to firms. We take this imperfect proxy because credit to households is not available for a long time span and for all the countries. However, we have checked that the effect of this variable is similar in a smaller sample of 13 countries where credit to households is available. We also control for real price changes of housing, and the variation of unemployment as a proxy of global macroeconomic conditions. We expect that growth of housing prices moves in the same direction as residential investment. As Poterba (1984) underlines, if any factors such as skilled construction workers, are in limited supply a rise in construction demand will increase the equilibrium price of houses. Moreover, the current growth of housing prices might be a proxy of anticipated price changes and households demand more housing when they expect a rise in housing prices.

C. *Results*

As expected, the growth rate of the population aged 20-49 is positively correlated with the residential investment rate. In Table 3, we take column (5) as our benchmark, since we do not have sufficient observations for credit in the case

[18] Most of the studies on the determinants of housing prices, however, use GDP per capita as a proxy for this variable (Takáts (2012), Bracke (2013)).

Table 2 – With 6 age groups

	Residential investment(% GDP)					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	5.78***	5.78***	5.83***	5.86***	5.91***	5.85***
	(0.65)	(0.66)	(0.66)	(0.67)	(0.70)	(0.68)
Population growth						
0 - 19	0.29	0.32	0.30	0.27	0.30	0.40*
	(0.21)	(0.21)	(0.21)	(0.21)	(0.22)	(0.21)
20 - 29	0.39***	0.39***	0.38***	0.36***	0.36***	0.41***
	(0.12)	(0.12)	(0.11)	(0.11)	(0.12)	(0.12)
30 - 49	0.72***	0.71***	0.68***	0.68***	0.68***	0.73***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.14)
50 - 64	0.02	0.00	−0.03	−0.03	−0.03	0.02
	(0.21)	(0.22)	(0.21)	(0.21)	(0.22)	(0.22)
65 - 74	−0.19	−0.20*	−0.20*	−0.19	−0.20	−0.16
	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)
75 plus	0.02	0.02	0.01	0.02	0.01	0.06
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.12)
Controls						
Real Disposable income growth		0.04	0.00	−0.00	−0.01	−0.04*
		(0.03)	(0.04)	(0.04)	(0.04)	(0.02)
Real House price growth			0.04**	0.04**	0.03**	0.03*
			(0.01)	(0.01)	(0.01)	(0.02)
Real long term interest rate				−0.03***	−0.03**	−0.03**
				(0.01)	(0.01)	(0.01)
ΔUnemployment					−0.08	−0.09
					(0.07)	(0.08)
Real Credit growth						0.02
						(0.03)
Observations	680	680	680	680	680	646
No. of countries	20	20	20	20	20	19
R^2 within	0.32	0.33	0.35	0.36	0.36	0.41

Note: Panel linear regression of the percentage share of GFCF dwelling on GDP, both in volumes, on the growth in percentage of the 0-19 years old, the 20-29 years old, the 30-49 years old,the 50-64 years old, the 65-74 years old and the 75 plus years old population, controlled by the growth in percentage of the index of real disposable income and the index of the real housing prices, the real long-term interest rates and the variation of unemployment, and the growth of real credit in percentage (which is not available for Luxembourg). The sample includes 20 OECD countries (Austria, Australia, Belgium, Canada, Switzerland, Denmark, Spain, Finland, France, UK, Greece, Ireland, Italy, Japan, Luxembourg, Netherlands, Norway, Portugal, Sweden, US) for the period 1980 to 2014. Regression includes country fixed-effect. We used clustered standard errors by country to correct for serial correlation and heteroskedasticity

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table 3 – Residential investment and population structure

	Residential investment(% GDP)					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	5.75***	5.73***	5.82***	5.91***	5.95***	5.88***
	(0.59)	(0.59)	(0.57)	(0.58)	(0.59)	(0.63)
Population growth						
0 - 19	0.20	0.21	0.20	0.19	0.21	0.30*
	(0.18)	(0.17)	(0.17)	(0.17)	(0.18)	(0.16)
20 - 49	1.04***	1.02***	0.98***	0.95***	0.97***	1.05***
	(0.20)	(0.21)	(0.19)	(0.19)	(0.20)	(0.21)
50 plus	−0.10	−0.11	−0.18	−0.19	−0.20	−0.08
	(0.36)	(0.36)	(0.35)	(0.35)	(0.35)	(0.38)
Controls						
Real Disposable income growth		0.03	−0.01	−0.01	−0.02	−0.06**
		(0.04)	(0.04)	(0.04)	(0.04)	(0.03)
Real House price growth			0.04**	0.04**	0.04**	0.03
			(0.02)	(0.01)	(0.01)	(0.02)
Real long term interest rate				−0.04***	−0.03***	−0.03***
				(0.01)	(0.01)	(0.01)
ΔUnemployment					−0.09	−0.09
					(0.07)	(0.07)
Real Credit Growth						0.04
						(0.03)
Observations	680	680	680	680	680	646
No. of countries	20	20	20	20	20	19
R^2 within	0.26	0.27	0.30	0.31	0.31	0.36

Note: Panel linear regression of the percentage share of GFCF dwelling on GDP, both in volumes, on the growth in percentage of the 0-19 years old, the 20-49 years old and the 50 ans plus years old population, controlled by the the growth in percentage of the index of real disposable income and the index of the real housing prices, the real long-term interest rates and the variation of unemployment, and finally the growth of real credit in percentage (which is not available for Luxembourg). The sample includes 20 OECD countries (Austria, Australia, Belgium, Canada, Switzerland, Denmark, Spain, Finland, France, UK, Greece, Ireland, Italy, Japan, Luxembourg, Netherlands, Norway, Portugal, Sweden, US) for the period 1980 to 2014. Regression includes country fixed-effect. We used clustered standard errors by country to correct for serial correlation and heteroskedasticity

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

of Luxembourg. In this case, an increase of one percentage point in the growth of the population aged 20-49 implies an increase of 0.97 percentage points of residential investment as a share of GDP, everything else equal. The impact of the 50 and above age group, although insignificant, is interesting. They would have a depressing effect on residential investment. In this age range, household are already formed and housing demand decreases. This interpretation is reinforced by the 6 age groups regression (cf Table 2) showing that the 65-74 years old are driving this negative effect. Note however that the mortality rate is higher for this age group, leading to more vacancies in the housing stock. Hence, this demand effect will be reinforced by a supply effect as well. Demand and supply are not as well identified by this variable as by the growth of the population aged 20-49, which isolates a pure demand effect.

Consistent with the findings of Glaeser, Gottlieb and Gyourko (2012) who find that dwellings permits are weakly correlated with interest rates in the US, we find a prominence of demographic variables over financial variables. The real interest rates have the expected sign although a rather small effect. When real long term interest rates increase by 1pp, the share of residential investment in GDP decreases by only 0.03 percentage points. The growth of real credit to the private non financial sector is insignificant. However, if we take real credit to households for the sub-sample of countries for which it is available, the credit growth is significant but as a rather small impact on residential investment (when household's real credit growth increases by one percent, residential investment increases by 0.07 pp cf. Table C4 in Appendix C). As expected, the real housing price growth has a positive impact on residential investment but it is rather small, when prices increase by 1%, the residential investment rate only increases by 0.04 percentage points. The variation of unemployement has the expected negative sign, when conditions tighten on the labor market, people would demand less housing, but the coefficient is insignificant. As for the disposable income, this variable would affect the ratio of residential investment to GDP only if going out of unemployment influences differently residential investment and other components of GDP, such as consumption.

The correlation between the growth of population aged between 20 and 49 and the residential investment rate is robust to the inclusion of different controls. The different controls could be determined at the same time as residential investment, hence there is potentially simultaneity and endogeneity that could bias their coefficients. However, we are not interested on their causal impact on residential investment but rather on the impact of demographics on residential investment. The bias of their coefficients is not transmitted over the coefficients of demographics variables, because they remain stable throughout the specifications. Moreover, the controls are weakly correlated with the age group growth[19] and do not present multicollinearity[20], which undermines a potential contamination between those different coefficients.

D. *Discussion*

Is the age profile effect on housing investment also observed on other components of GDP? Table 4 show evidence of the specificity of the cycle of residential investment, compared to consumption and non-residential investment. Growth rates of consumption and nonresidential investment, as well as growth rates of the ratio of consumption and non-residential investment to GDP, are positive during the periods of negative growth of both residential investment and residential investment to GDP. Interestingly, we also observe an average decline of the consumption to GDP ratio when the investment to GDP ratio increases. It suggests potential substitution between the two main components of household expenditures. Thus, these simple stylized facts are evidence that the demand for housing differs from the demand for consumption goods, even if these two types of goods are mainly demanded by households. Moreover, we estimate the same equation for consumption and non-residential investment as a share of GDP, we do not observe the same age distribution pattern as the one observed for residential investment (cf. Table C5 in Appendix C).

One can argue that investment in housing could be substituable with other form of savings. Indeed, following this reasoning, the recent decrease of residential investment over GDP could be driven by an increase in aggregate savings. This increase could be link to an increase of precautionary savings by the 20-49 years old, who expect lower retirement pensions in the future. We check whether this mechanism holds in our sample. In Table C6, we observe that saving as a share of GDP is positively correlated with residential investment ratio, which refutes a potential substituability between the two variables. Moreover, the coefficient of the impact of the 20-49 years old population on the residential investment ratio is robust to the inclusion of saving as a control.

In the previous estimations and argumentation, we have not taken into account the household size. Household size is however an obvious and well

[19] For example, the correlation between the age group growth of the population aged 20-49 and the controls does not exceed 0.15 (except for the growth of credit where the correlation is equal to 0.25).

[20] The mean VIF between all the explanatory variables is 1.32. The rule of thumb is to have a VIF under 10.

documented determinant of housing demand (Börsch-Supan (1986), Glaeser, Gyourko and Saks (2006), Caldera and Johansson (2013)): for a constant number of people, the housing demand is higher if the household size is smaller (or in other words, if the number of households increases). In fact, the negative relationship between household size and residential investment is straightforward because, by definition, the number of primary residences equals the number of households. Macroeconomic cross-country studies have not controlled for the household size (Bracke (2013), Lindh and Malmberg (2008), Takáts (2012)) because annual data are not available at the country level. For this same reason, we have not used it as a control in the previous estimations.

Variation of the household size can be due to two main factors. The first factor is a decrease in the fertility rate. This is not an issue in our specification because we control for the growth of children (0-19 age group). The second factor is that more people are living alone because of sociological reasons (divorce rate, increase of enrollment in university, increase in life expectancy etc.).

Our estimations would suffer from an omitted variable bias if, indeed, household size and the growth rate of the population aged 20-49 were systematically negatively correlated. However, we have shown that the change of an age group is volatile and varied largely within a short time span. On the contrary, the decreasing trend of the household size is a structural phenomenon in advanced countries since the 1980s, that is not related to the growth of the age group 20-49, as shown on Figure B9. Overall, Figure B9 does not show a negative correlation between household size and the growth of the population aged 20-49 for the years where data is available. We also see, on this figure, that the early 2000s boom in housing construction, in most OCDE countries, did not correspond to a period of exceptional decrease in household size, whereas it did correspond to a period of unusual increase in the growth of the age group 20-49 in most countries. Hence, we do not find evidence that the unability to control for the household size is creating an omitted variable bias in our specification.

We have found evidence of a robust correlation between the growth of the population aged 20-49 years and the residential investment rate. We now turn our attention to whether this correlation can be identified as causal.

Table 4 – Cycle statistics

	Growth of the residential investment(%GDP)	Decline of the residential investment(%GDP)
Residential investment		
Growth	8.56	−5.68
Contribution to GDP growth	0.49	−0.31
Consumption		
Growth	2.61	1.72
Δ of the share in GDP	−0.24	0.05
Non residential investment		
Growth	3.59	2.10
Δ of the share in GDP	0.12	0.03
GDP growth	2.95	1.66
Growth of population		
0-19	−0.32	−0.40
20-49	0.65	0.48
50 plus	1.48	1.42
Controls		
Real House Price growth	4.47	−0.65
Personal Disposable income growth	2.01	1.46
Real Long term interest rate	1.10	1.64
ΔUnemployment	−0.15	0.32
Real Credit Growth	5.29	3.54
Real Household Credit Growth	6.70	3.70
Number of observations	300	400

Note: Mean value of country-year pooled data for the 20 OECD countries between 1980 and 2014. All variables are expressed in percentage or percentage points.

III. Population Age Structure and Housing Investment : an IV Approach

A. *IV Strategy and Reduced Form*

If people tend to immigrate to countries with a more dynamic housing sector and higher GDP growth, or residents emigrate when a country experiences low growth (which can be due or correlated to a housing crisis), then it will be difficult to assess the causality between the growth of an age group and residential investment. We cannot disentangle if people are moving because there is more new dwellings available, or more job opportunities correlated with a booming economy, or whether it is their arrival that leads to an increase in the demand for housing, and to a higher residential investment rate. This problem is likely to be important for our estimations since a major share of migrants are between 20 and 50 years old.[21] If there is such an endogenous mechanism between migrations and the real investment rate, the coefficient of the growth of population aged 20-49 will be biased upward. To our best knowledge, there is no paper attempting to cope with this issue. Furthermore, since there is no published statistics on the age of the flow of migrants (see Section IV), it is not straightforward to see whether this potential endogeneity problem is sizeable and whether the cycle of the growth of population aged 20-49 (as see on Figure 1) is strongly influenced by migrations.

To address and estimate this potential endogeneity bias, we use an instrumental variable approach. Without migrations and with a constant mortality for each age group over time, the growth of the population aged 20-49 today will be approximatively equal to the growth of the population aged 0-29, 20 years ago. Moreover, the evolution of an age group 20 years ago is unlikely to be correlated with the current residential investment rate.

This method allows us to both discuss whether the potential endogeneity problem is sizeable (that is if migrations contribute significantly to the growth rate of the age group) and to provide a causal examination.

[21] According to the wallchart of the United Nations Department of Economic and Social Affairs on the Age and Sex of Migrants in 2011, the median age of the international migrant stock is 39 years old. It varies from 29 to 49 years old depending of the degree of development of the country of origin. Age distribution of migrants increases until 39 years old and then decreases. Nevertheless, this study is only about the stock, i.e. foreign-born or migrants of foreign nationality residing in the host country, thus the age of the stock is higher than the age at the time of arrival. Furthermore, this statistics does not include migrations, in and out of a country, by domestic residents (typically students studying abroad and coming back to work in their home country).

The first stage of our instrumental variable approach is:

$$(1) \qquad \Delta Pop\,20\text{-}49_{c,t} = \alpha_0 + \alpha_1 \Delta Pop\,0\text{-}29_{c,t-20} + \alpha_2 X_{c,t} + \gamma_c + \xi_{c,t}$$

where $\Delta Pop\,20 - 49_{c,t}$ is the growth rate in percentage of the population aged between 20 and 49 years old at time t in country c, and $\Delta Pop\,0 - 29_{c,t-20}$ is the growth rate in percentage of the population aged between 0 and 29 years old at time $t - 20$ in country c. Equation (1) still includes the conventional controls ($X_{c,t}$) and country fixed effects (γ_c). The second stage is:

$$(2) \qquad \begin{aligned} GFCFdwellings/GDP_{c,t} &= \alpha_0 + \alpha_1 \Delta Pop0\text{-}19_{c,t} + \alpha_2 \widehat{\Delta Pop20\text{-}49}_{c,t} \\ &+ \alpha_3 \Delta Pop50+_{c,t} + \alpha_4 X_{c,t} + \gamma_c + \xi_{c,t} \end{aligned}$$

and the reduced form:

$$(3) \qquad \begin{aligned} GFCFdwellings/GDP_{c,t} &= \alpha_0 + \alpha_1 \Delta Pop0\text{-}19_{c,t} + \alpha_2 \Delta Pop0\text{-}29_{c,t-20} \\ &+ \alpha_3 \Delta Pop50+_{c,t} + \alpha_4 X_{c,t} + \gamma_c + \xi_{c,t} \end{aligned}$$

Figure 2a and Figure 2b show that the potential endogeneity problem can be very important for some countries in the sample, where the actual growth rate of the population aged 20-49 would have been very different without migration in the 20 previous years. Note that we use the term "migration" as the difference between immigration and emigration (that is "net migration"). Increased labor mobility between countries will not increase "migration" if the number of immigrants compensate the number of emigrants. The Spanish and Irish cases in the early 2000s are the most striking (Figure 2a): in these countries, the growth rate of population aged 20-49 between 2000 and 2006 would have been on average -0.13% and 0.29% respectively without migrations (during the 20 previous years), whereas the actual average growth rates reached 1.72% in Spain and 2.5% in Ireland. Another interesting information from Figure 2a and Figure 2b is that the countries of low migration are not significantly different from countries of high migration. In particular, we do not observe a "Schengen effet" in Europe that would have caused a high increase in net migrations in the Schengen area countries only beginning in the 1990s. Indeed, the migration boom of the 2000s was not predictable, the 1998 projection of the United Nations were unable to forecast it (cf. Figure B8).

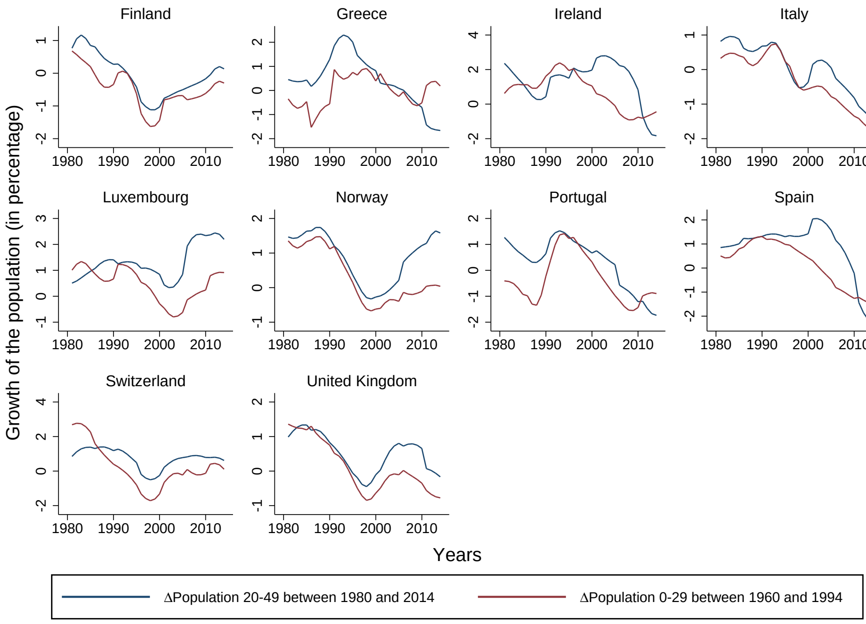
As a consequence, there is no evidence that a common shock would have affected a part of the sample at the same time. There is also no evidence that countries systematically differ in their level of migrations. As seen on Figure 2a, there are migration booms in some countries concentrated in a small part of the sample. Thus, the level of migrations is more likely to be related to country-specific characteritics which varied overtime. We interpret this result as further evidence of a potential link between migrations and the business (and housing) cycle.

The previous discussion highlights an important characteristic of our instrument. The past age group growth will be a strong instrument except for countries that had experienced substantial net migration. The instrument is expected to be weak in countries where, on average, the actual growth of the 20-49 significantly differs from the growth rate of population aged 0-29, twenty years before. To overcome this difficulty, we divide the sample in two. In a first subsample, we group the countries where the difference between the growth rate of population 20-49 in year t and the growth rate of the population aged 0-29 in year $t-20$, is below average. This subsample is called "Low migration group" (Figure 2b) whereas the other subsample is called "High migration group" (Figure 2a). The rationale for simply dividing the sample in two subsamples with 10 countries in each is twofold. First, as seen on Figure 2a, there is a sufficient number of countries in our sample where the instrument is expected to be strong. Second, dividing the sample in half (rather than 8 vs. 12, for example) ensures comparability of degrees of freedom between subsamples. We will use this comparison in the next sections to study further the magnitude of the potential endogeneity bias. The high population flow sample includes Switzerland, Spain, Finland, UK, Greece, Ireland, Italy, Luxembourg, Norway and Portugal. The low population flow sample includes Austria, Australia, Belgium, Canada, Denmark, France, Japan, Netherlands, Sweden and the US.

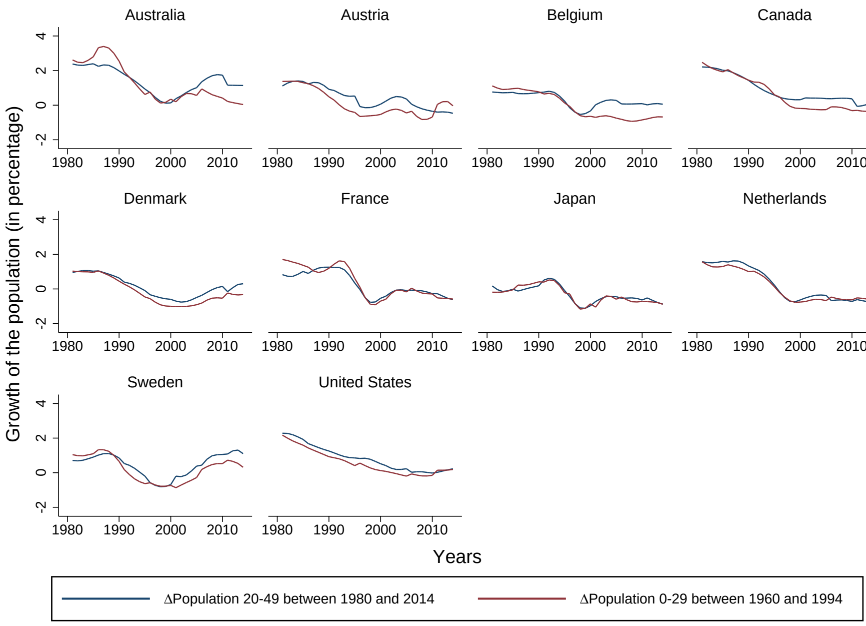
B. *Results*

We first report the results of the reduced form estimation (Table 5).[22] The reduced form is directly interpretable and interesting in itself since it measures the causal impact of the predicted growth of population aged 20-49 under the assumptions of zero net migrations and a constant mortality rate by age group. In the sample of countries that experienced low population flows, the predicted growth of the 20-49 years old (without migrations) would be enough to explain

[22] We do not control for credit in these regressions since the correlation observed in the previous section was 0 (and insignificant) and since we lack data for the Luxembourg.



A. High Flow Group



B. Low Flow Group

Figure 2. Comparison of the 20-49 – 20 Years Apart

Table 5 – Reduced form

	(1) All sample	(2) Low population flow	(3) High population flow
Constant	6.38***	5.07***	7.29***
	(0.63)	(1.14)	(0.28)
Population growth between 1960 and 1994			
0 - 29	0.50**	0.92**	0.22
	(0.24)	(0.40)	(0.21)
Population growth between 1980 and 2014			
0-19	0.18	0.34	0.13
	(0.18)	(0.20)	(0.25)
50 plus	−0.22	0.26	−0.55**
	(0.40)	(0.63)	(0.19)
Controls			
Real Disposable income growth	−0.01	0.04	−0.02
	(0.04)	(0.02)	(0.05)
Real House price growth	0.05**	0.03	0.06**
	(0.02)	(0.03)	(0.02)
Real long term interest rate	−0.05***	−0.03	−0.05***
	(0.01)	(0.06)	(0.01)
ΔUnemployment	−0.03	0.00	−0.07
	(0.10)	(0.06)	(0.13)
Observations	680	340	340
No. of countries	20	10	10
R^2 within	0.16	0.22	0.17

Note: Panel linear regression of the percentage share of GFCF dwelling on GDP, both in volumes, on the the growth in percentage of the 0-29 years old between 1960 and 1994, controlled by the growth in percentage of the 0-19 years old and the 50 years old and above between 1980 and 2014, the index of real disposable income and the index of the real housing prices, the real long-term interest rates and the variation of unemployment. The high population flow sample includes Switzerland, Spain, Finland, Greece, UK, Italy, Ireland, Luxembourg, Norway and Portugal while the low population flow sample contains the rest of the sample. Regression includes country fixed-effect. We used clustered standard errors by country to correct for serial correlation and heteroskedasticity

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table 6 – Instrumental variable

	(1) All sample	(2) Low population flow	(3) High population flow
Population growth			
0-19	0.12	0.34**	0.10
	(0.16)	(0.16)	(0.21)
20 - 49	0.65**	1.31***	0.29
	(0.29)	(0.41)	(0.24)
50 plus	−0.30	0.38	−0.64***
	(0.37)	(0.55)	(0.18)
Controls			
Real Disposable income growth	−0.01	0.00	−0.02
	(0.04)	(0.03)	(0.04)
Real House price growth	0.04**	0.02	0.06**
	(0.02)	(0.03)	(0.02)
Real long term interest rate	−0.04***	−0.04	−0.04***
	(0.01)	(0.05)	(0.01)
ΔUnemployment	−0.06	−0.07	−0.08
	(0.05)	(0.07)	(0.08)
Montiel-Pflueger Robust weak IV test	No	No	Yes
K-P F stat	64.83	62.05	23.89
Underidentification test	0.00	0.03	0.01
Observations	680	340	340
No. of countries	20	10	10
R^2 within	0.29	0.17	0.28

Note: Second stage panel linear regression of an IV specification of the percentage share of GFCF dwelling on GDP, both in volumes on the growth in percentage of the 0-19 years old, the 20-49 years old and the 50 ans plus years old population, controlled by the growth in percentage of the index of real disposable income and the index of the real housing prices, the real long-term interest rates and the variation of unemployment. The under-identification test reports the p-value that uses the LM and Wald versions of the Kleibergen and Paap (2006) rk statistic, which are cluster-robust statistics. For the weak instrument test, we report Kleibergen-Paap Wald rk F statistic, and the results of the Olea and Pflueger (2013) test ("Yes" meaning that the instrument is weak), both statistics are robust to serially correlated and hetereskedastic disturbances. The high population flow sample includes Switzerland, Spain, Finland, Greece, UK, Italy, Ireland, Luxembourg, Norway and Portugal, the low population flow sample contains the rest of the sample. Regression includes country fixed-effect. We used clustered standard errors by country to correct for serial correlation and heteroskedasticity
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

the current variation of the residential investment. Indeed, we find that an increase of 1 pp of the predicted growth of the population aged 20-49 (proxied by the growth rate of the population aged 0-29, twenty years before) would increase the residential investment rate by 0.92 pp. However, in the other sub-sample, the predicted growth of the 20-49 age group is not sufficient to explain residential investment dynamics. The coefficient is not significant. In this second subsample, migrations contribute so much to the growth dynamics of population aged 20-49 that it is impossible to find an effect on the residential investment rate when using past demographic data. Results over the whole sample still show a positive and significant effect of the changes in the age profile on residential investment, but the coefficient is lower than in the estimations on the "low migration sample". Without taking into account the potential weakness of our instrument, we would have found a lower impact, although still significant.

Table 6 displays the results of the IV regression. The Montiel-Plfueger test assesses, as expected, that the instrument is strong in the "Low migration sample" but weak in the "High migration sample". It also points to a strong instrument for the whole sample.[23] The IV regression confirms the result of the reduced form. In the low population flow sample, there is a significant positive effect of the growth rate of the population aged 20-49 on the residential investment rate. The coefficient is slightly higher than in the reduced form case because the first stage coefficient are less than one (cf Table C7). When the growth rate of the population aged 20-49 in year t increased by 1pp, the ratio of residential investment to GDP in year t increases by around 1.3 pp. When estimated over the whole sample, this coefficient is still significative but lower (0.65) than with the "Low migration" sample. As with the reduced form, the impact is not signicative in countries of high migration.

An additional important feature of the reduced form and IV estimations is that the effects of house prices and real interest rates disappear when the past demographic growth is a strong instrument. Demographic changes appear to be a better predictor of the cycle of the residential investment rate than any other macroeconomic variable.

It is interesting to have a look at our initial regression without instrumentation to assess the size and direction of the bias (Table 7). The size of the bias actually looks rather small, the coefficient of the population aged 20-49 is 0.93

[23] The Montiel-Plfueger test is the robust equivalent of the standard F-test. It should be used when the estimation includes robust standard errors, as it takes into account heteroskedasticity and serial correlation of the residuals (Olea and Pflueger (2013)).

in the low flow group versus 1.10 in the high flow group. Nevertheless, in the high flow group, the real long-term interest rate and the variation of unemployment have a significant effect. These correlations might indicate that people are moving in and out of the country for economic and financial reasons such as the state of the labor market and the level of interest rates. Indeed, Ferreira, Gyourko and Tracy (2010) show that negative equity and rising interest rates have an impact on the mobility of owners at the city level. In the high migration group, people move more and the growth of the 20-49 is more correlated to residential investment than in countries with low population flow. Note that, in the case of a weak instrument, the demand effect linked to the arrival of new residents could be reinforced by a supply effect if a part of population flows are constituted by people working in construction (this supply effect might diminish labor costs and thus the cost of construction).

Furthermore, as expected from the reduced form and IV results described previously, the growth of the population aged 20-49 catches all the explanatory power of the cycle of the residential investment rate in countries that didn't experience high migration. In those countries, it would have been possible to forecast the actual growth of the 20-49, and, accordingly, to foreseen the path of the residential investment rate.

Table 7 – Residential investment and sub-samples

	(1) All sample	(2) Low population flow	(3) High population flow
Constant	5.95***	5.09***	6.86***
	(0.59)	(1.34)	(0.33)
Population growth			
0 - 19	0.21	0.21	0.32
	(0.18)	(0.16)	(0.23)
20 - 49	0.97***	0.93**	1.10***
	(0.20)	(0.40)	(0.20)
50 plus	−0.20	0.14	−0.61**
	(0.35)	(0.69)	(0.25)
Controls			
Real Disposable income growth	−0.02	0.01	−0.04
	(0.04)	(0.03)	(0.05)
Real House price growth	0.04**	0.03	0.04*
	(0.01)	(0.02)	(0.02)
Real long term interest rate	−0.03***	−0.04	−0.03***
	(0.01)	(0.06)	(0.01)
ΔUnemployment	−0.09	−0.03	−0.17**
	(0.07)	(0.04)	(0.07)
Observations	680	340	340
No. of countries	20	10	10
R^2 within	0.31	0.19	0.42

Note: Panel linear regression of the percentage share of GFCF dwelling on GDP, both in volumes, on the growth in percentage of the 0-19 years old, the 20-49 years old and the 50 ans plus years old population, controlled by the growth in percentage of the index of real disposable income and the index of the real housing prices, the real long-term interest rates and the variation of unemployment. The high population flow sample includes Switzerland, Spain, Finland, Greece, Ireland, Luxembourg, Norway and Portugal for the period 1980 to 2014, the low population sample contains the rest of the sample. Regression includes country fixed-effect. We used clustered standard errors by country to correct for serial correlation and heteroskedasticity

* *p < 0.1,* ** *p < 0.05,* *** *p < 0.01.* Standard errors in parentheses.

IV. More on population flows

A. *Data : reconstruction of migration balance by age groups*

For some countries with high migration in our sample, the age groups 20 years apart are very different. We want to investigate further whether, in this group, residential investment is correlated with migration flows of people aged 20-49. By migration, we mean populations flows in and out of a country whatever the nationality of the people concerned. We still won't be able to identify a causal effect of migrations flows because they could be endogenous to residential investment. However, we can compare the size of the coefficient of migration flows and the size of the coefficient of natural balance, in order to assess whether the endogeneity bias is likely to be large.

No statistical institutes report migration flows by age groups. They usually estimates global migration balance annually without age decomposition. The UN made available the international migrant stock by age groups for a few census (1990, 2000, 2010 and 2013), but the stock recovers only people of foreign nationality or foreign-born, rather than migration flows.

To estimate net migration by age, we use the forward method of the cohort component population method, a commonly used model in demographics (cf. Smith, Tayman and Swanson (2013)). Usually, this method is employed to reconstruct migration flows for 5 years interval. However for our sample, we can find population and survival rates for each age annually, which therefore allow us to reconstruct net migration annually.

We can reconstruct the migration balance between $t-1$ and t for each age x using the probability of surviving from one age to the next and the population for each age x.

$$Netmigr^{x}_{t-1/t} = Pop^{x}_{t} - S^{x}_{x-1,t-1} Pop^{x-1}_{t-1}$$

where $Netmigr^{x}_{t-1/t}$ is the net migration of person of age x, Pop^{x}_{t} is the population of age x at time t, $S^{x}_{x-1,t-1}$is the probability of surviving from age x to age $x+1$ at time $t-1$, and $Pop^{x-1}_{t-1}$is the population of age $x-1$ at time $t-1$. $S^{x}_{x-1,t-1} Pop^{x-1}_{t-1}$ represents the expected population of age x at time t, substracting this expected population to the actual population of age x at time t , assuming that the difference is due to migration, this relationship gives us the number of net migrants. The forward method estimates the number of net migrants at the end of the period and assumes that all deaths are to non-migrants. However, since we are applying this method for one year to the next, this assumption would create a small bias.

This method assumes that population changes, which are not due to fertility and mortality, are due to migrations. Nevertheless, these differences could also be related to migration errors in the census counts or boundary changes from one census period to the next. There are also differences in measure, some countries have a resident registery such as Austria, Denmark, Finland, Norway, Netherlands, Italy, Luxembourg and Japan whereas others do not force their residents to register. Moreover, statistical institutes when estimating population for each age make assumptions about migrations that we recover by applying our method.

To circumvent those potential approximations, we take 3 year moving average of our variables, and focus on the net migration of the 20-49 years old. We also control by the natural variation of the population (which is the actual variation of the population minus the net migration balance estimated before). Due to data availability, we can only apply our method starting in 1986. Using 3 years moving averages reduces our estimation sample to 1988- 2013.

B. *Specification and results*

We used the migration data that we reconstructed as another mean to confirm the results obtained by the instrumental variable regression. Indeed, the natural variation of the population could be thought as an exogenous phenomenon since it is only related to fertility behavior from 20 to 49 years ago (for the age range of interest : the population aged between 20 and 49 years old). In our IV approach, we capture long-term trends of migration throughout the period. Here, we look at short-term migration dynamics and their link with residential investment cycle. Moreover, our data reconstruction did implicitly took into account the mortality variation (through the survival probabilities) which was not taken into account in the IV strategy. We decompose the 20-49 years old between their variation related to the net migration and the one due to the natural balance. Natural balance of an age group is the difference between the age group variation and the net migration. We deflate by the population aged between 20 and 49 years old, in order to have coefficients comparable to the age group growth of the other age ranges (the 0-19 and the 50 years old and above). We keep the same controls as before, and control for the dynamics of the other age groups, leading to the following specification :

$$\begin{aligned} GFCFdwellings/GDP_{c,t} &= \alpha_0 + \alpha_1 \frac{NetMigr20\text{-}49_{c,t}}{Pop20\text{-}49_{c,t-1}} + \alpha_2 \frac{NatBal20-49_{c,t}}{Pop20\text{-}49_{c,t-1}} \\ &+ \alpha_3 \Delta Pop0\text{-}19_{c,t} + \alpha_4 \Delta Pop50+_{c,t} + \alpha_5 X_{c,t} + \gamma_c + \xi_{c,t} \end{aligned}$$

where $\frac{NetMigr20\text{-}49_{c,t}}{Pop20\text{-}49_{c,t-1}}$ represents the contribution of the net migration of the 20-49 to the growth of the 20-49, $NatBal20\text{-}49_{c,t} = \Delta Pop20\text{-}49_{c,t} - NetMigr20\text{-}49_{c,t}$ is the natural balance for the 20-49 age group and $\frac{NatBal20-49_{c,t}}{Pop20\text{-}49_{c,t-1}}$ is the contribution of the natural balance of the 20-49 to the growth of the 20-49. $\Delta Pop0\text{-}19_{c,t}$ and $\Delta Pop50+_{c,t}$ are the growth rates of the 0-19 and the 50 years old and above respectively. As before, $X_{c,t}$ and γ_c are the time-varying control and the country fixed-effects.

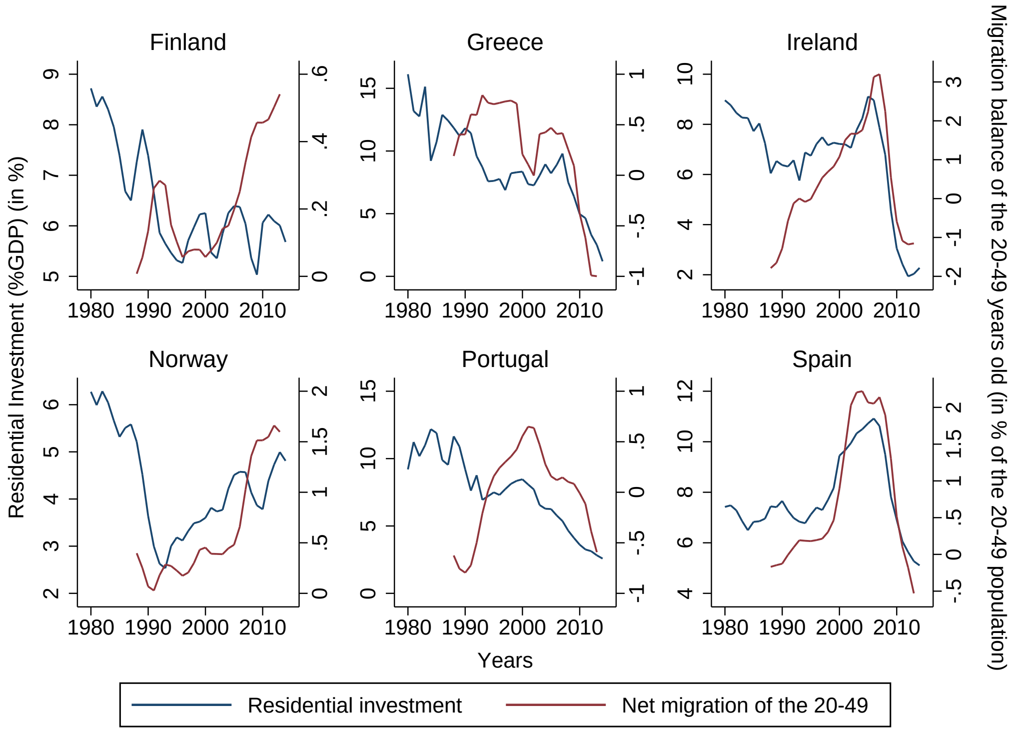


FIGURE 3. NET MIGRATION OF THE 20-49 AND RESIDENTIAL INVESTMENT

According to Table 8, the impact of newly arrived immigrants (in the last 3 years) is stronger than the impact of the growth rate of the natural balance of people aged 20-49, in countries of high migration. In effect, if the contribution of the net migration to the the growth of the 20-49 years old increase by one percent, residential investment increase by 1.08 pp as a share of GDP in the countries of high population flows whereas the natural balance increase residential investment by 0.75 pp. This result tends to support the claim that the coefficient is indeed biased upward because of endogeneity, although the bias looks rather small. It supports the need for a causal identification as used in this paper. However, another interpretation is possible if immigrants have different characteristics from residents. In this case, the coefficient would be biased upward not because of endogeneity but because of different characteristics of the population. Reality probably lies in between.

The results confirm what we found in Section III. In the countries that did not experience substantial net migration flows, residential investment cycle is explained by the dynamics of the natural balance and not by the one of

Table 8 – Residential investment, migration and natural balances

	(1) All sample	(2) Low population flow	(3) High population flow
Constant	5.25***	4.07***	6.59***
	(0.63)	(0.62)	(0.42)
Migration balance			
20-49	0.62**	−0.46	1.08***
	(0.27)	(0.37)	(0.18)
Natural balance			
20-49	1.07***	1.62***	0.75**
	(0.27)	(0.36)	(0.24)
Population growth			
0 - 19	0.46**	0.84**	0.44*
	(0.19)	(0.28)	(0.23)
50 plus	0.20	1.08**	−0.72**
	(0.45)	(0.36)	(0.31)
Controls			
Real House price growth	0.04***	0.05**	0.03*
	(0.01)	(0.02)	(0.02)
Real Disposable income growth	0.09***	0.06*	0.11***
	(0.02)	(0.03)	(0.02)
Real long term interest rate	−0.03***	−0.03	−0.03***
	(0.01)	(0.06)	(0.00)
ΔUnemployment	−0.05	0.10	−0.16***
	(0.07)	(0.12)	(0.05)
Observations	520	260	260
No. of countries	20	10	10
R^2 within	0.43	0.41	0.57

Note: Panel linear regression of the percentage share of GFCF dwelling on GDP, both in volumes, on the 3 years moving average of the migration balance of the 20-49 years old, the natural balance (which is equal to the variation of the population minus the migration balance) for the same age range, controlled by the growth rate in percentage of the 0-19 years old, the 50 and above years old population, the index of real disposable income and the index of the real housing prices, the real long-term interest rates and the variation of unemployment. The high population flow sample includes Switzerland, Spain, Finland, Greece, Ireland, Luxembourg, Norway and Portugal for the period 1980 to 2014, the low population sample contains the rest of the sample. Regression includes country fixed-effect. We used clustered standard errors by country to correct for serial correlation and heteroskedasticity

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

migration flows. An increase of one percent of the contribution of natural balance to the growth of the population increases residential investment by 1.62 pp, which is of a rather similar magnitude to the 1.3 pp impact that we found with our instrumental variable approach.

V. Conclusion

The residential investment rate is well explained by the growth rate of the population aged 20-49. Since the latter is actually highly cyclical, it is not surprising that residential investment is much more volatile than GDP: the variation of the 20-49 age group affects directly the demand for housing but not the other components of GDP. Since demographic projections point to negative growth of the population aged 20-49 in OECD countries in the several next decades, we should expect that the growth of housing investment will be slower than the growth of GDP. This could however be reversed if new migration flows changed demographic projections. Since the contribution of residential investment to GDP growth is actually low (0.02 pp on average in our sample), the slowdown of housing construction in percentage of total output will not necessarily lead to a prolonged stagnation. However, further research is needed on the potential multiplier effect of housing investment on the other components of GDP, in order to assess the potential aggregate deflationary effect of low housing growth. Beyond long term discussions, a better understanding of the housing cycle is of primary interest for fiscal, monetary and macroprudential policies. When discussing housing construction booms and busts, relative to GDP, policymakers should closely keep an eye on the demographic cycle. Financial variables and monetary policy, on the contrary, appear to play at most a minor role.

GFCF in dwellings For most of the countries of the sample, we use the Gross Fixed Capital Formation in dwellings, in billions of national currency, 2010 constant prices, not seasonally adjusted , made publicly available by the European Commission in its AMECO database. For Canada and Switzerland respectively, we take the same data but from the OECD Annual (respectively Quaterly) National Accounts because of longer time availability. Those series are in millions, we convert it to billions. For the Australia series, we are missing the 2014 observation, we take it from the OECD Quaterly National Accounts in 2012-2013 chained price.

GDP For all the countries, we take the total Gross Domestic Product at market prices, in national currency, 2010 constant prices from the AMECO database. To have the residential investment as a share of GDP (in percentage), we take the ratio of the GFCF in dwellings on the GDP and multiply it by 100.

Population by Five-Year Age Groups To construct our age groups, we take the Annual Population by five-year age groups for both sexes from the United Nations World Population Prospect 2015. It is the population as of 1 July of the year indicated classified by five-year age groups (0-4, 5-9, 10-14, ..., 95-99, 100+). Data are presented in thousands. We sum the five-year age groups to obtain our 3 age groups : the 0-19, the 20-49 and the 50 and above. To obtain the age group growth in percentage, we take the variation of the log multiplied by 100. For the age group for people 20 years ago 20 years younger, we take the five-year age groups from 1960 to 1994 and sum them to obtain the 0-29 and the 30 and above.

Real long term interest rates

Nominal long term interest rates For most of the countries, we take the Long-Term Interest Rate On Government Bonds, seasonally adjusted, in percentage from the OECD Economic Outlook. For Greece, Ireland and Luxembourg, we take the nominal long term interest rates, not seasonally adjusted, from the AMECO database. For Greece, there is missing data for the years 1989 to 1991, we interpolate by using a Kriging method of 1.5.

Inflation We use the consumer price for all items in percentage change from previous period provided by the OECD statistics. To obtain the real long term interest rates (r), we use the Fisher equation : $1 + r = \frac{1+i}{1+\pi}$ where$i$ is the nominal long term interest rate and π is the inflation.

Real Personal Disposable Income We used the harmonized and consolidated data base made by the Federal Reserve of Dallas. The Personal Disposable Income series are quoted in per capita terms using working-age population (population aged between 15 and 64 years old). Real values are computed using the personal consumption expenditure deflator. The series are indexes where 2005=100. For exhaustive details on this database, see Mack and Martínez-García (2011). For Austria and Portugal, we take the personal disposable income provided by Oxford Economics, in billions of current euros, seasonally adjusted for Portugal and not for Austria. We divide them by the population between 15 and 64 years old (constructed from the UN WPP described above) and deflated the ratio by the Personal Consumption Expenditure deflator taken from OECD statistics (we deflate by the log of the PCE for Portugal and Greece to avoid the distortion related to periods of hyper inflation). For Greece, we take the same approach but we use the net national disposable income (because the personal disposable income is not available) in billions of current euros, not seasonally adjusted, from AMECO. We rebased all the variable to 2005=100.

Real Housing Prices We used the harmonized and consolidated data base made by the Federal Reserve of Dallas. The series are indexes where 2005=100. Real values are computed using the personal consumption expenditure deflator. For a comprehensive description of the database, see Mack and Martínez-García (2011). For Austria, Norway, Switzerland, Sweden, Denmark, Portugal, Greece , United Kingdom, Australia, Italy, Luxembourg, the price refer to the price of new and existing dwellings, whereas for Belgium, Canada, France, Spain, Ireland, Finland, United States, Netherlands, the price is only for existing dwellings. For Austria, Greece and Portugal, we do not have a consistent price series throughout the period. To proxy the price we use the deflator of the GFCF in dwellings from AMECO and we rebase it to have 2005=100.

Unemployment For all the countries, we take the unemployment rate in percentage of civilian labour force, according to the Eurostat definition, from the AMECO database.

Real Credit We take the credit to private non-financial sector from all sectors, at market value in domestic currency, adjusted for breaks from the BIS (Bank of International Settlements). They capture the outstanding amount of credit at the end of the reference quarter. Credit is provided by domestic banks, all other sectors of the economy and non-residents. In terms of financial instruments, credit covers the core debt, defined as loans, debt securities, currency and deposits. For credit to households, we take the credit to households and

NPISHs from all sectors, at market value in domestic currency, adjusted for breaks from the BIS. To obtain the real credit, we deflate by the Consumer Price Index (2010=100) from the OECD.

Population for each age To obtain the population for each age, for most of the European countries, we use the population on first of January by age and sex in Eurostat. For France, we use the same variable for metropolitan France provided by INSEE. For Australia, US, Canada and Japan, we use Total population (both sexes combined) by single age, major area, region and country, annually for 1950-2100 from the UN World Population Prospect 2015.

Survival probability For most of the European countries, we use the probability of surviving between exact ages from the life tables provided by Eurostat. For France, the probability for 2013 is missing, we complete it by the mortality table of 2013 provided by INSEE. For the UK, we complete the missing values from 1986 to 1992 by the life tables provided by the Office of National Statistics (averaging the probability of dying between to exact ages for men and women, and taking one minus this expression to have the survival probability). For Australia, we compute the survival probability between to exact age using historical mortality rates provided by the Australian Government Actuary. For Canada, we use the life tables available from the Canadian Human Mortality database. For Japan, we use life table made available by the National Institute of Population and Social Security Research. For the US, we use the life tables from the Human Mortality database and the center for Disease Control and Prevention.

Appendix B Figures

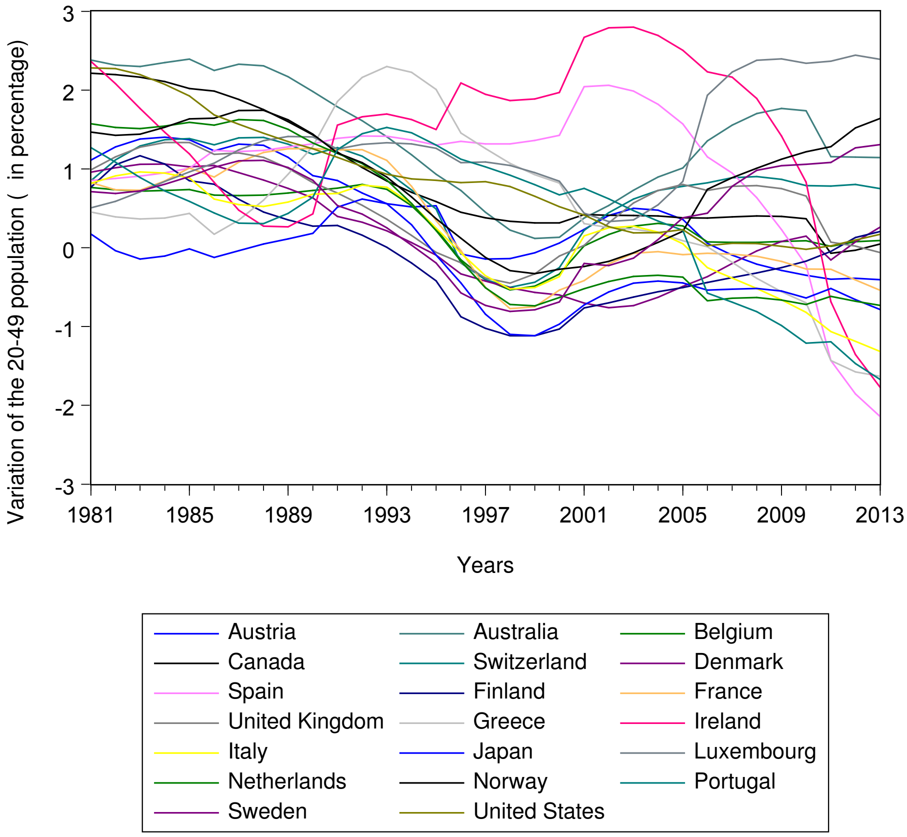


Figure B1. Growth of the 20-49 in all the countries

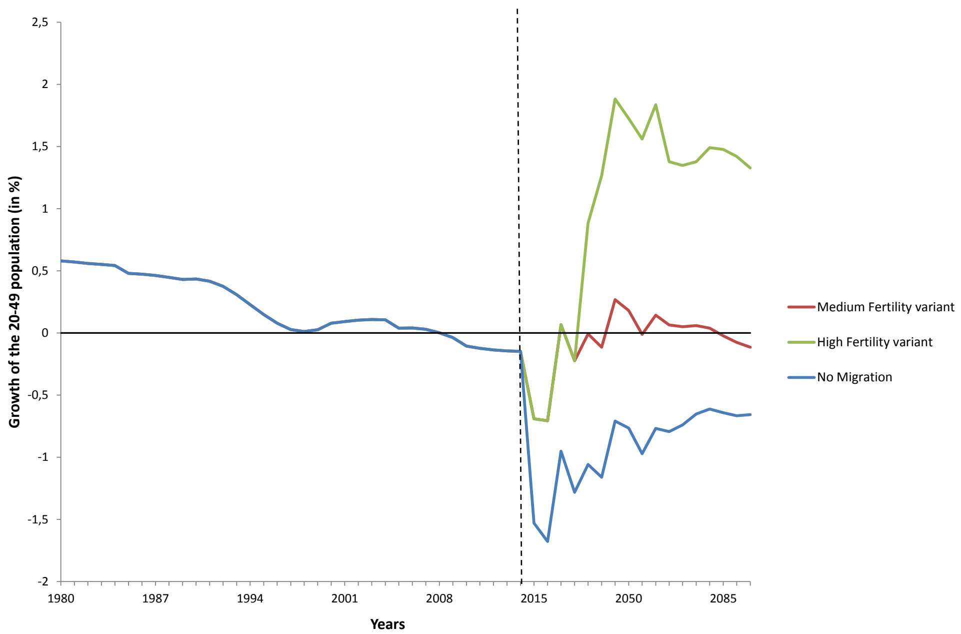
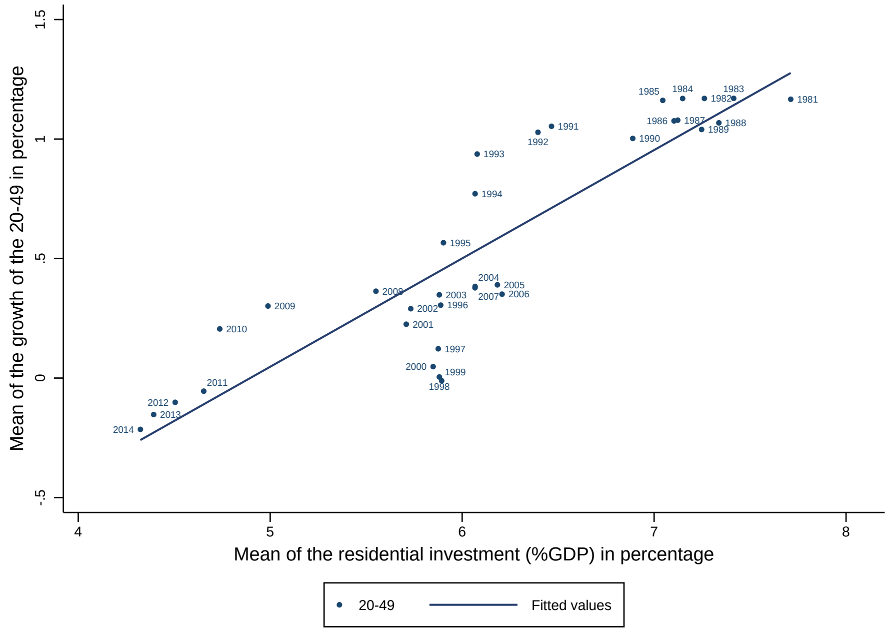
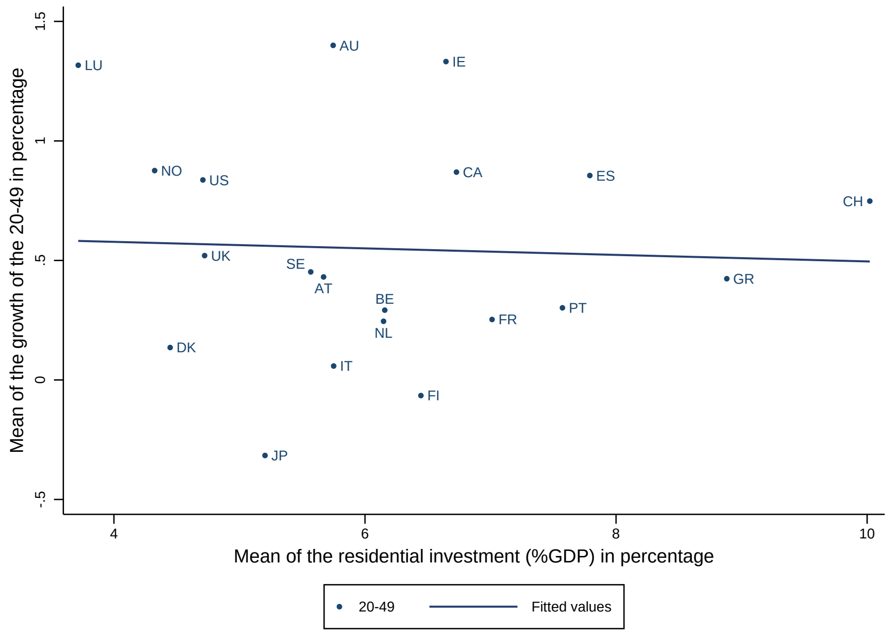


Figure B2. Forecast of the 20-49 from 2015 to 2100



A. Mean of All the Countries for Each Year



B. Mean of All the Years for Each Country

Figure B3. Scatterplot Between vs Within

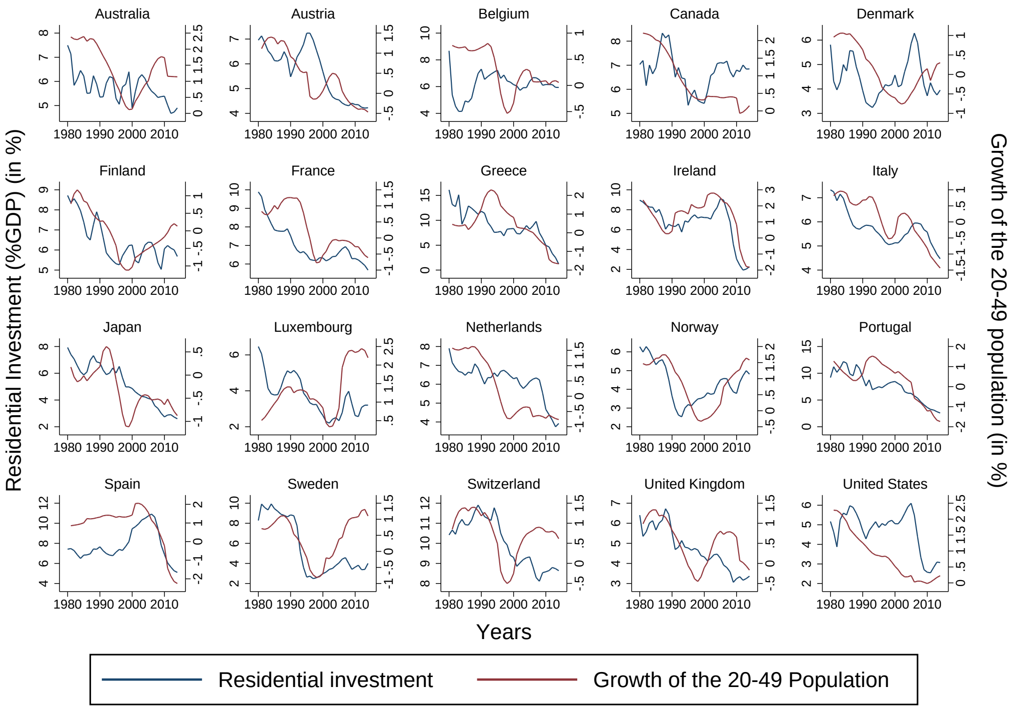


FIGURE B4. RESIDENTIAL INVESTMENT (%GDP) AND GROWTH OF THE 20-49

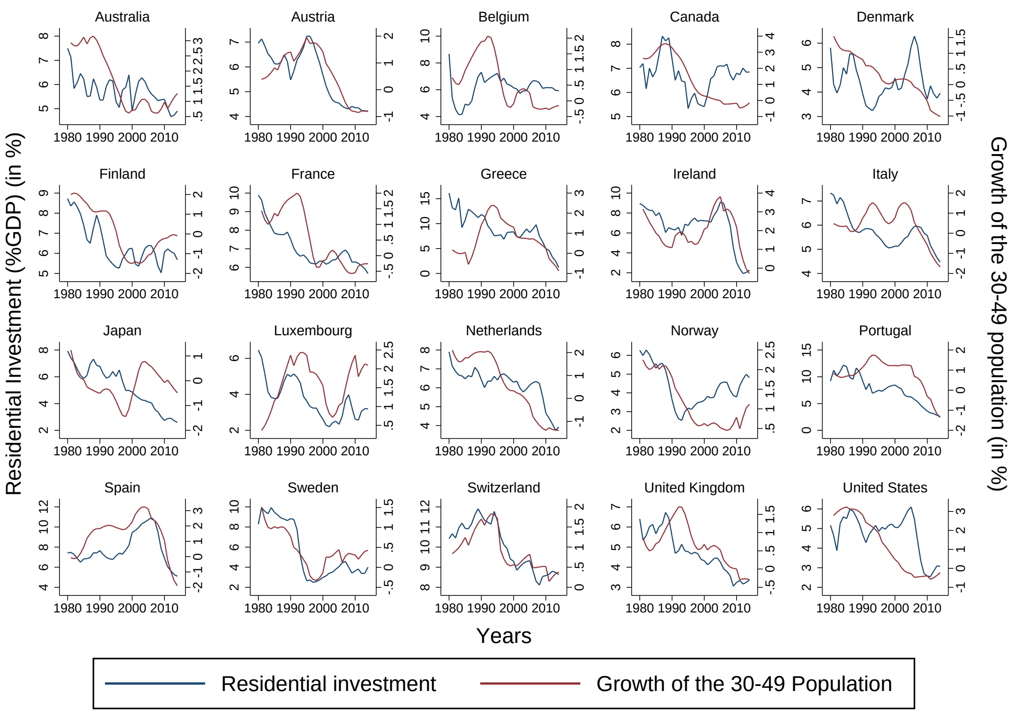


FIGURE B5. RESIDENTIAL INVESTMENT (%GDP) AND GROWTH OF THE 30-49

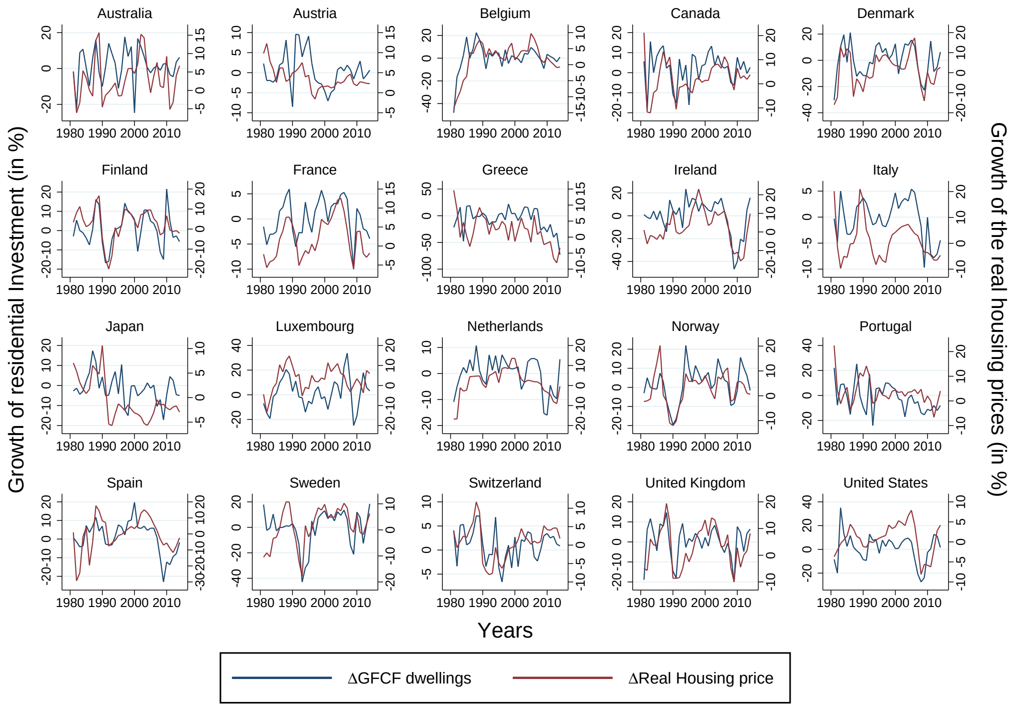


FIGURE B6. GROWTH OF RESIDENTIAL INVESTMENT VS HOUSING PRICES

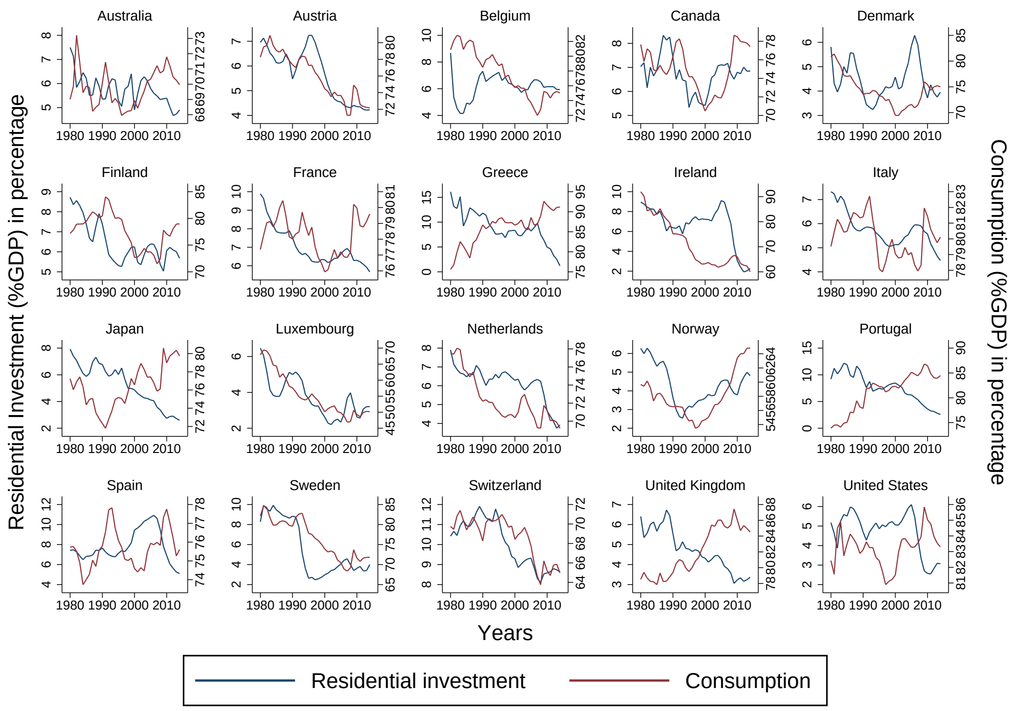


FIGURE B7. RESIDENTIAL INVESTMENT AND CONSUMPTION (%GDP)

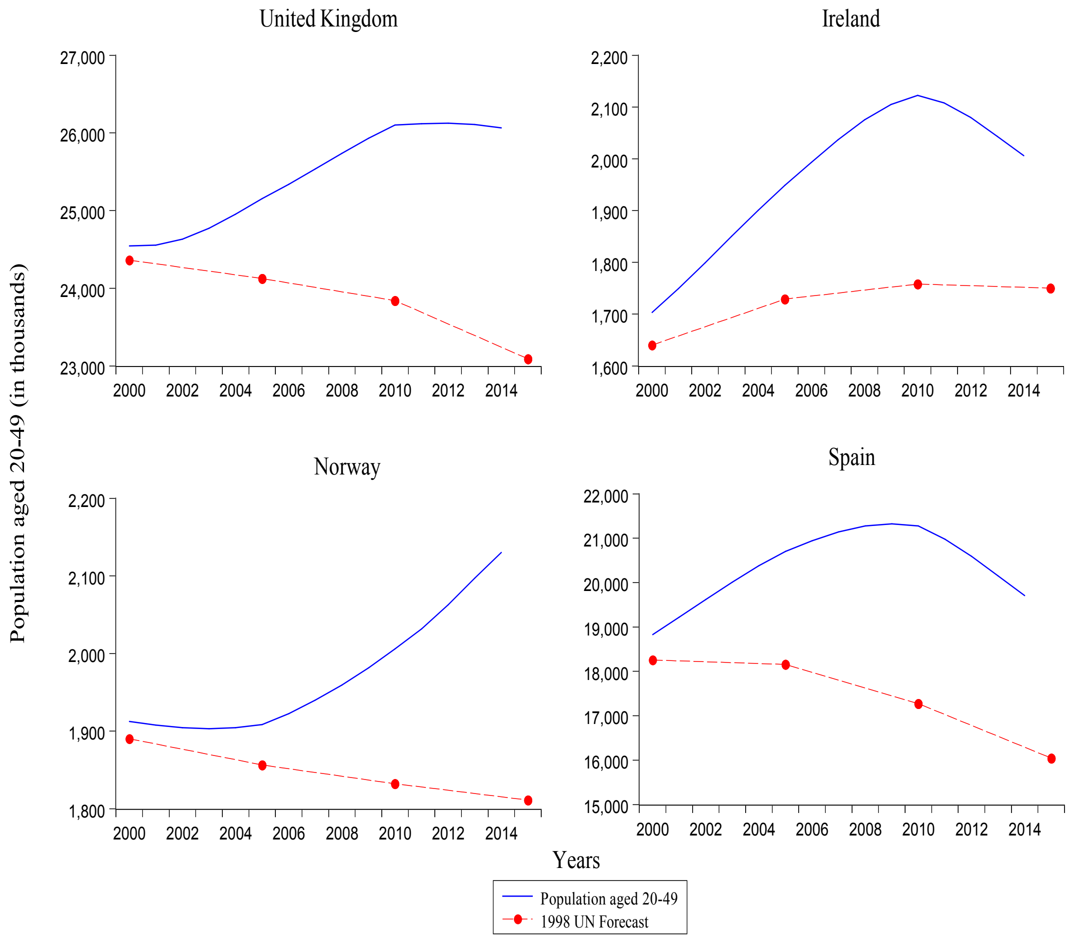


Figure B8. 20-49 population and its forecast in 1998

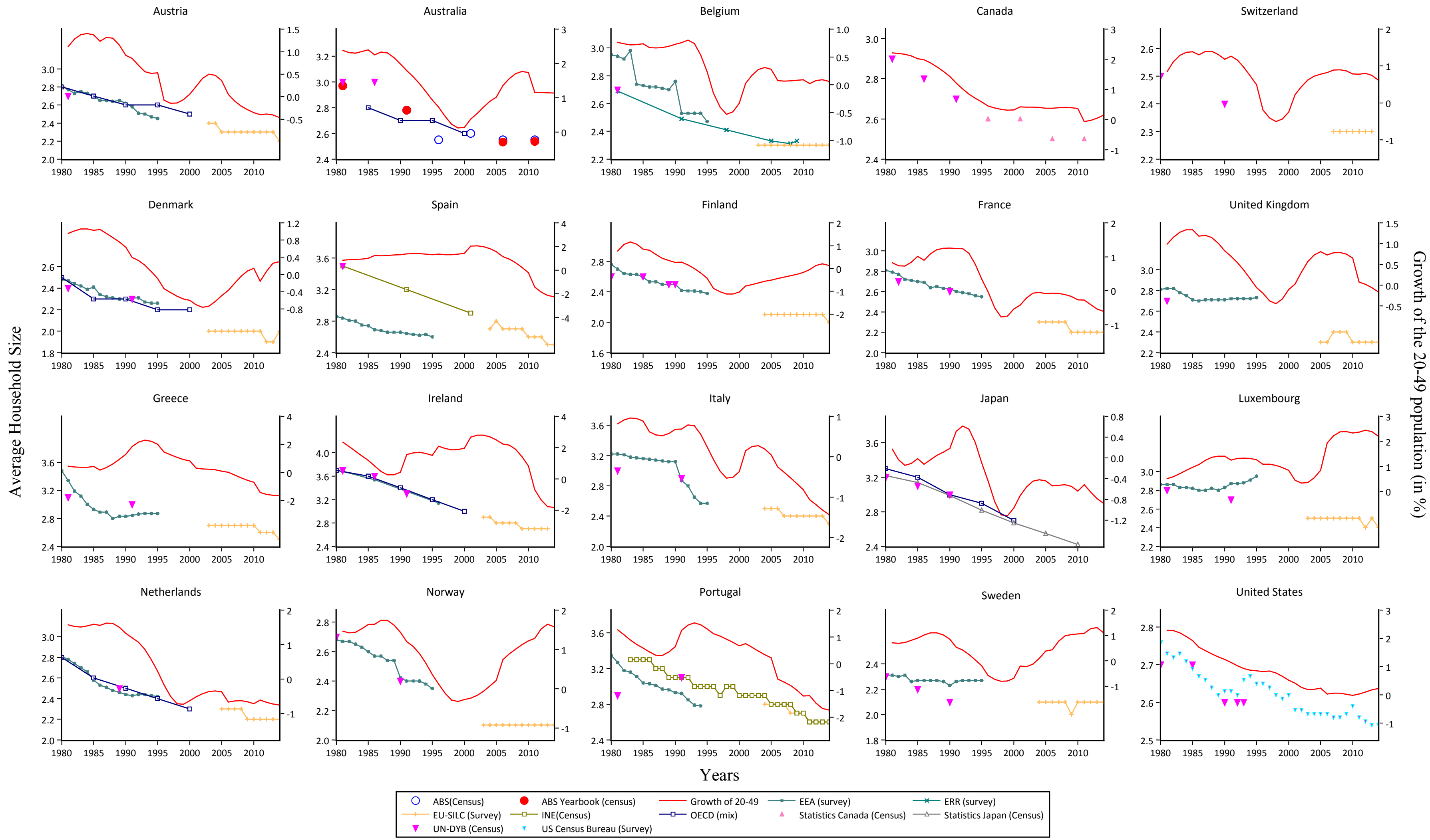


FIGURE B9. AVERAGE HOUSEHOLD SIZE AND GROWTH OF THE 20-49 POPULATION

Appendix C Tables

Table C1 – Various tests

	RE vs OLS Breusch-Pagan LM	FE vs RE Sargan-Hansen statistic	Heteroskedasticity Modified Wald test	Autocorrelation Wooldridge test
p-value	0.00	0.00	0.00	0.00

Note: The null hypothesis in the Breusch-Pagan LM test is that variances across entities is zero i.e. that there is random effects.For the robust Hausman test, the null hypotesis is that random effect is appropriate. For the modified Wald test, the null is homoskedasticity. For the Woolridge test, the null is no serial correlation.

Table C2 – Cross-section depedence test

	Test for cross-section dependence	
	Breusch-Pagan independence test	Pasaran CD test
p-value	0.00	0.00

Note: H_0: no cross-sectional depedence.

We use second generation panel unit root tests since our panel presents evidence of cross-section depedence (Table C2). The Levin-Lin-Chu and Harris-Tzavalis tests are based on regression t statistics that are subsequently adjusted to reflect the fact that under the null hypothesis, the t statistics have a nonzero mean because of the inclusion of panel-specific means or trends. The Breitung (2000) test takes a different approach, transforming the data before computing the regressions so that the standard t statistics can be used. The Breitung test assumes that all panels have a common autoregressive parameter. The null hypothesis is that all series contain a unit root. The alternative hypothesis is that $\rho < 1$ so that the series are stationary. Breitung and Das (2005) remark that the test also has power in the heterogeneous case, where each panel is allowed to have its own autoregressive parameter, though the test is optimal in the case where all panels have the same autoregressive parameter. The Breitung (2000) Monte Carlo simulations suggest that his test is substantially more powerful than other panel unit-root tests for the modest-size dataset he considered (N=20, T=30), it is why we use it here.

The Pesaran (2007) CIPS test allows for allows for heterogeneity in the autoregressive coefficient of the Dickey-Fuller regression and allows for the presence of a single unobserved common factor with heterogeneous factor loadings in the data. The statistic is constructed from the results of panel-member-specific (A)DF regressions where cross-section averages of the dependent and

Table C3 – Panel Unit Root tests

	P-values	
	Breitung	Pesaran (2007)
Main variables		
Residential investment	0,23	0,93
Residential Investment (%GDP)	0,04	0,00
Population 0-19	1,00	0,04
Population 20-49	1,00	1,00
Population 50 plus	1,00	0,47
Controls		
Real disposable Income	1,00	0,10
Real House Price	0,99	0,41
Real Long Term Interest Rate	0,00	0,16
Unemployment	0,06	0,13
Real Credit	0,99	0,01
Real Household Credit	0,98	0,83

Note: H_0: all series contain a unit root.

independent variables (including the lagged differences to account for serial correlation) are included in the model (referred to as CADF regressions). We choose the number of lags by applying the BIC criterion. Under the null of nonstationarity the test statistic has a non-standard distribution.

Table C4 – Credit to Households

	(1) Benchmark	(2) Credit	(3) Household's Credit
Constant	5.88***	5.13***	5.08***
	(0.63)	(0.86)	(0.84)
Population growth			
0 - 19	0.30*	0.12	0.16
	(0.16)	(0.22)	(0.21)
20 - 49	1.05***	1.02***	0.95***
	(0.21)	(0.30)	(0.26)
50 plus	−0.08	0.18	0.15
	(0.38)	(0.46)	(0.43)
Controls			
Real Disposable income growth	−0.06**	−0.02	−0.04
	(0.03)	(0.02)	(0.03)
Real House price growth	0.03	0.02	0.01
	(0.02)	(0.02)	(0.02)
Real long term interest rate	−0.03***	−0.09	−0.08
	(0.01)	(0.07)	(0.06)
ΔUnemployment	−0.09	−0.03	−0.01
	(0.07)	(0.08)	(0.08)
Real Credit Growth	0.04	0.05	
	(0.03)	(0.04)	
Real Household Credit Growth			0.07**
			(0.03)
Observations	646	442	429
No. of countries	19	13	13
R^2 within	0.36	0.34	0.36

Note: Panel linear regression of the percentage share of GFCF dwelling on GDP, both in volumes, on the growth in percentage of the 0-19 years old, the 20-49 years old, and the 50 years old and above population, controlled by the growth in percentage of the index of real disposable income and the index of the real housing prices, the real long-term interest rates and the variation of unemployment. In column (1), we control for the growth of real credit to the private non-financial sector for all the sample (except Luxembourg because of data availability). Column (2) represents the same regression except for the subsample for which credit to households is available. Finally, in column (3), we control for credit to households. The sample includes 20 OECD countries (Austria, Australia, Belgium, Canada, Switzerland, Denmark, Spain, Finland, France, UK, Greece, Ireland, Italy, Japan, Luxembourg, Netherlands, Norway, Portugal, Sweden, US) for the period 1980 to 2014. Regression includes country fixed-effect. We used clustered standard errors by country to correct for serial correlation and heteroskedasticity

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table C5 – Macroeconomic aggregates and population structure

	Residential Investment	Non residential investment	Consumption
	(1)	(2)	(3)
Constant	5.88***	13.46***	80.04***
	(0.63)	(0.68)	(1.87)
Population growth			
0 - 19	0.30*	0.33	−0.11
	(0.16)	(0.29)	(0.35)
20 - 49	1.05***	−0.37	−0.37
	(0.21)	(0.25)	(0.51)
50 plus	−0.08	0.76*	−2.73**
	(0.38)	(0.42)	(1.21)
Controls			
Real Disposable income growth	−0.06**	0.10**	0.00
	(0.03)	(0.04)	(0.07)
Real House price growth	0.03	−0.00	−0.04
	(0.02)	(0.02)	(0.04)
Real long term interest rate	−0.03***	−0.01	0.00
	(0.01)	(0.01)	(0.02)
ΔUnemployment	−0.09	0.15**	0.50***
	(0.07)	(0.07)	(0.17)
Real Credit growth	0.04	0.16***	−0.05
	(0.03)	(0.05)	(0.07)
Observations	646	646	646
No. of countries	19	19	19
R^2 within	0.36	0.27	0.29

Note: Panel linear regression of the percentage share of GFCF (resp. Consumption, non residential GFCF) dwelling on GDP, both in volumes, on the growth in percentage of the 0-19 years old, the 20-49 years old and the 50 years old and above population, controlled by the growth in percentage of the index of real disposable income and the index of the real housing prices, the real credit, the level of the real long-term interest rates and the variation of unemployment. The sample includes 19 OECD countries (Austria, Australia, Belgium, Canada, Switzerland, Denmark, Spain, Finland, France, UK, Greece, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Sweden, US) for the period 1980 to 2014. Regression includes country fixed-effect. We used clustered standard errors by country to correct for serial correlation and heteroskedasticity

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table C6 – Savings

	Benchmark	Reduced form		
	(1)	(2)	(3)	(4)
	All sample	All sample	Low population flow	High population flow
Constant	5.65***	5.85***	4.79***	6.74***
	(0.60)	(0.67)	(1.19)	(0.38)
Population growth between 1960 and 1994				
0-29		0.50*	0.93*	0.24
		(0.26)	(0.42)	(0.20)
Population growth between 1980 and 2014				
0-19	0.19	0.20	0.31	0.23
	(0.19)	(0.19)	(0.18)	(0.30)
20 - 49	0.82***			
	(0.16)			
50 plus	−0.33	−0.34	0.26	−0.79***
	(0.37)	(0.44)	(0.63)	(0.23)
Controls				
Real Disposable income growth	−0.02	−0.02	0.03	−0.02
	(0.03)	(0.03)	(0.03)	(0.03)
Real House price growth	0.02	0.03	0.03	0.03
	(0.01)	(0.02)	(0.03)	(0.02)
Real long term interest rate	−0.03**	−0.03***	−0.03	−0.03***
	(0.01)	(0.01)	(0.05)	(0.01)
ΔUnemployment	−0.02	0.04	0.02	0.01
	(0.06)	(0.09)	(0.06)	(0.13)
Savings (%GDP)	0.11***	0.14***	0.06	0.15***
	(0.02)	(0.03)	(0.06)	(0.03)
Observations	679	679	339	340
No. of countries	20	20	10	10
R^2 within	0.40	0.30	0.23	0.40

Note: Panel linear regression of the percentage share of GFCF dwelling on GDP, both in volumes, on the growth in percentage of the 0-19 years old, the 20-49 years old or the 0-29 years old between 1960 and 1994, the 50 years old and above population, controlled by the growth in percentage of the index of real disposable income and the index of the real housing prices, the real credit, the level of the real long-term interest rates and the variation of unemployment, and the share, in percentage of net national savings in GDP. The sample includes 20 OECD countries (Austria, Australia, Belgium, Canada, Switzerland, Denmark, Spain, Finland, France, UK, Greece, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Sweden, US, Luxembourg) for the period 1980 to 2014. Regression includes country fixed-effect. We used clustered standard errors by country to correct for serial correlation and heteroskedasticity
* *p < 0.1*, ** *p < 0.05*, *** *p < 0.01*. Standard errors in parentheses.

Table C7 – First stage

	Growth of the 20-49 between 1980 and 2014		
	(1)	(2)	(3)
	All sample	Low population flow	High population flow
Population growth between 1960 and 1994			
0-29	0.77***	0.70***	0.77***
	(0.10)	(0.09)	(0.16)
Population growth between 1980 and 2014			
0-19	0.09*	−0.00	0.12**
	(0.05)	(0.06)	(0.05)
50 plus	0.12	−0.09	0.29**
	(0.12)	(0.10)	(0.15)
Controls			
Real Disposable income growth	0.01	0.03	0.01
	(0.01)	(0.02)	(0.01)
Real House price growth	0.01***	0.00	0.02***
	(0.01)	(0.00)	(0.01)
Real long term interest rate	−0.01***	0.00	−0.01***
	(0.00)	(0.01)	(0.00)
ΔUnemployment	0.04	0.06***	0.05
	(0.07)	(0.02)	(0.11)
Observations	680	340	340
No. of countries	20	10	10
Partial R^2	0.40	0.62	0.31

Note: First stage panel linear regression of the growth in percentage of the 20-49 years old between 1980 and 2014 on the growth in percentage of the 0-29 years old between 1960 and 2014, controlled by the growth in percentage of the 0-19 years old and the 50 years old and above between 1980 and 2014, the growth of the index of real disposable income and the index of the real housing prices, the real long-term interest rates and the variation of unemployment.The high population flow sample includes Switzerland, Spain, Finland, Greece, Ireland, Luxembourg, Norway and Portugal for the period 1980 to 2014, the low population sample contains the rest of the sample. Regression includes country fixed-effect. We used clustered standard errors by country to correct for serial correlation and heteroskedasticity

* *p < 0.1,* ** *p < 0.05,* *** *p < 0.01.* Standard errors in parentheses.

REFERENCES

Agnello, Luca, and Ludger Schuknecht. 2011. "Booms and busts in housing markets: Determinants and implications." *Journal of Housing Economics*, 20(3): 171–190.

Bénétrix, Agustín S., Barry Eichengreen, and Kevin H. O'Rourke. 2012. "How housing slumps end." *Economic Policy*, 27(72): 647–692.

Benhabib, Jess, Richard Rogerson, and Randall Wright. 1991. "Homework in Macroeconomics: Household Production and Aggregate Fluctuations." *Journal of Political Economy*, 99(6): 1166–87.

Börsch-Supan, Axel. 1986. "Household formation, housing prices, and public policy impacts." *Journal of Public Economics*, 30(2): 145–164.

Bracke, Philippe. 2013. "How long do housing cycles last? A duration analysis for 19 OECD countries." *Journal of Housing Economics*, 22(3): 213–230.

Breitung, Jörg. 2000. "The Local Power of Some Unit Root Tests for Panel Data."

Breitung, Jörg, and Samarjit Das. 2005. "Panel unit root tests under cross–sectional dependence." *Statistica Neerlandica*, 59(4): 414–433.

Caldera, Aida, and Åsa Johansson. 2013. "The price responsiveness of housing supply in OECD countries." *Journal of Housing Economics*, 22(3): 231–249.

Davis, Morris A., and Jonathan Heathcote. 2005. "Housing And The Business Cycle." *International Economic Review*, 46(3): 751–784.

Eichholtz, Piet, and Thies Lindenthal. 2014. "Demographics, human capital, and the demand for housing." *Journal of Housing Economics*, 26(C): 19–32.

Engelhardt, Gary V., and James M. Poterba. 1991. "House prices and demographic change: Canadian evidence." *Regional Science and Urban Economics*, 21(4): 539–546.

Fair, Ray C., and Kathryn M. Dominguez. 1991. "Effects of the Changing U.S. Age Distribution on Macroeconomic Equations." *American Economic Review*, 81(5): 1276–94.

Faubert, Violaine, Éric Monnet, and Camille Sutter. 2015. "Despite the recovery of purchasing power, housing construction should keep falling in 2015." INSEE Notes de conjoncture.

Favara, Giovanni, and Jean Imbs. 2015. "Credit Supply and the Price of Housing." *American Economic Review*, 105(3): 958–92.

Ferreira, Fernando, Joseph Gyourko, and Joseph Tracy. 2010. "Housing busts and household mobility." *Journal of Urban Economics*, 68(1): 34–45.

Glaeser, Edward L., and Joseph Gyourko. 2005. "Urban Decline and Durable Housing." *Journal of Political Economy*, 113(2): 345–375.

Glaeser, Edward L., Joseph Gyourko, and Raven E. Saks. 2006. "Urban growth and housing supply." *Journal of Economic Geography*, 6(1): 71–89.

Glaeser, Edward L., Joshua D. Gottlieb, and Joseph Gyourko. 2012. "Can Cheap Credit Explain the Housing Boom?" In *Housing and the Financial Crisis. NBER Chapters*, 301–359. National Bureau of Economic Research, Inc.

Green, Richard, and Patric H. Hendershott. 1996. "Age, housing demand, and real house prices." *Regional Science and Urban Economics*, 26(5): 465–480.

Higgins, Matthew. 1998. "Demography, National Savings, and International Capital Flows." *International Economic Review*, 39(2): 343–69.

Holland, Steven A. 1991. "The baby boom and the housing market: Another look at the evidence." *Regional Science and Urban Economics*, 21(4): 565–571.

Iacoviello, Matteo, and Stefano Neri. 2010. "Housing Market Spillovers: Evidence from an Estimated DSGE Model." *American Economic Journal: Macroeconomics*, 2(2): 125–64.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2015. "Betting the house." *Journal of International Economics*, 96(S1): S2–S18.

Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti. 2015. "Credit Supply and the Housing Boom." National Bureau of Economic Research, Inc NBER Working Papers 20874.

Kleibergen, Frank, and Richard Paap. 2006. "Generalized reduced rank tests using the singular value decomposition." *Journal of Econometrics*, 133(1): 97–126.

Knoll, Katharina, Moritz Schularick, and Thomas Steger. 2014. "No Price Like Home: Global House Prices, 1870-2012." C.E.P.R. Discussion Papers CEPR Discussion Papers 10166.

Kydland, Finn E., Peter Rupert, and Roman Sustek. 2012. "Housing Dynamics over the Business Cycle." National Bureau of Economic Research, Inc NBER Working Papers 18432.

Leamer, Edward E. 2007. "Housing is the business cycle." *Proceedings – Economic Policy Symposium – Jackson Hole*, 149–233.

Leamer, Edward E. 2015. "Housing Really Is the Business Cycle: What Survives the Lessons of 2008–09?" *Journal of Money, Credit and Banking*, 47(S1): 43–50.

Lee, Gabriel, Philipp Schmidt-Dengler, Bernhard Felderer, and Christian Helmenstein. 2001. "Austrian Demography and Housing Demand: Is There a Connection." *Empirica*, 28(3): 259–276.

Lindh, Thomas, and Bo Malmberg. 2008. "Demography and housing demand – what can we learn from residential construction data?" *Journal of Population Economics*, 21(3): 521–539.

Mack, Adrienne, and Enrique Martínez-García. 2011. "A cross–country quarterly database of real house prices: a methodological note." *Globalization and Monetary Policy Institute Working Paper*, 99.

Mankiw, Gregory N., and David N. Weil. 1989. "The baby boom, the baby bust, and the housing market." *Regional Science and Urban Economics*, 19(2): 235–258.

Muellbauer, John, and Anthony Murphy. 1997. "Booms and Busts in the UK Housing Market." *Economic Journal*, 107(445): 1701–27.

Ohtake, Fumio, and Mototsugu Shintani. 1996. "The effect of demographics on the Japanese housing market." *Regional Science and Urban Economics*, 26(2): 189–201.

Olea, José Luis Montiel, and Carolin Pflueger. 2013. "A Robust Test for Weak Instruments." *Journal of Business & Economic Statistics*, 31(3): 358–369.

Peek, Joe, and James A. Wilcox. 1991. "The Measurement and Determinants of Single–Family House Prices." *Real Estate Economics*, 19(3): 353–382.

Pesaran, M. Hashem. 2007. "A simple panel unit root test in the presence of cross–section dependence." *Journal of Applied Econometrics*, 22(2): 265–312.

Poterba, James M. 1984. "Tax Subsidies to Owner–occupied Housing: An Asset–Market Approach." *The Quarterly Journal of Economics*, 99(4): 729–52.

Shiller, Robert J. 2007. "Low Interest Rates and High Asset Prices: An Interpretation in Terms of Changing Popular Economic Models." *Brookings Papers on Economic Activity*, 38(2): 111–134.

Smith, Stanley K., Jeff Tayman, and David A. Swanson. 2013. "Implementing the Cohort–Component Method." In *A Practitioner's Guide to State and Local Population Projections*. 155–183. Springer.

Takáts, Előd. 2012. "Aging and house prices." *Journal of Housing Economics*, 21(2): 131–141.

Topel, Robert, and Sherwin Rosen. 1988. "Housing Investment in the United States." *Journal of Political Economy*, 96(4): 718–740.