DOCUMENT
DE TRAVAIL
N° 591

DEMOGRAPHIC CYCLE, MIGRATION AND HOUSING INVESTMENT: A CAUSAL EXAMINATION

Eric Monnet and Clara Wolf

April 2016

BANQUE DE FRANCE

EUROSYSTÈME

DIRECTION GÉNÉRALE DES ÉTUDES ET DES RELATIONS INTERNATIONALES

DEMOGRAPHIC CYCLE, MIGRATION AND HOUSING INVESTMENT: A CAUSAL EXAMINATION

Eric Monnet and Clara Wolf

April 2016

Demographic Cycle, Migration and Housing Investment : a Causal Examination

By Eric Monnet and Clara Wolf*

*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 especially grateful to Yannick Kalantzis for many comments and suggestions at the early stage of this project. We also thank Philippe Bracke, Guillaume Chapelle, Nicolas Coeurdacier, Violaine Faubert, Thomas Grjebine, Jacques Friggit, Guy Lévy-Rueff, Philippe Martin, John Muellbauer, Pierre Sicsic, Vichett Oung as well as seminar participants at Banque de France, Sciences Po, T2M conference, Oxford Symposium on Population, Migration and Environment. The paper was written when Clara Wolf was visiting the Banque de France.

Résumé

Nous étudions l'évolution de l'investissement résidentiel en proportion du PIB, depuis 1980 pour 20 pays de l'OCDE, et montrons qu'elle est fortement corrélée à la croissance de la population âgée de 20-49 ans. Nous développons ensuite une nouvelle méthode pour identifier l'effet causal de la croissance de cette tranche d'âge. Utilisant des données démographiques passées comme instrument afin de contourner le possible biais d'endogénéité entre les migrations et le cycle du logement, nous trouvons qu'une hausse d'1% de la population âgée de 20-49 ans augmente l'investissement résidentiel en proportion du PIB de 1.3 points de pourcentage. Les variations démographiques prédisent mieux l'évolution du taux d'investissement résidentiel que n'importe quelle autre variable financière ou macroéconomique prise en compte dans nos estimations. (*JEL* E32, J11, R21)
Mots clés : cycle, logement, démographie, migrations

Abstract

We study residential investment over GDP in 20 OECD countries since 1980, and show that it is closely associated with the growth dynamics of population aged 20-49. We develop a new method to uncover the causal effect of the growth of the 20-49 age group. Using past demographic data as an instrument to avoid potential endogeneity between migration and the housing cycle, we find that a 1% increase in the population aged 20-49 increases the residential investment rate by 1.3 pp. Demographic changes are a better predictor of the residential investment rate than any macroeconomic or financial variable we control for. (*JEL* E32, J11, R21)
Keywords: business cycle, housing, demography, migration

Non technical summary

In France, as well as in the Eurozone, housing investment in 2015 is 25% below its 2007 level. This fall has been more pronounced than for any other component of GDP. Hence, the ratio of housing investment to GDP in the Eurozone has fallen from 6,8% in 2007 to 4,8% in 2015 (6,9 % to 5,4 % in France). This reflects that the growth of housing investment is higher than the growth of GDP during periods of output expansion, but is lower during downturns.

Previous research had already shown that housing investment is one of 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 consumption and housing investment are both mainly produced by households. For this reason, Edward Leamer famously entitled his Jackson Hole 2007 paper "Housing IS the business cycle".

So, why does housing investment vary so much, compared to the other components of GDP? What drives the cycle of the housing investment rate? Previous discussions have emphasized the role of financial conditions or supply constraints. We argue instead that demography is the main factor explaining the peculiar cyclicality of housing investment. Looking at a sample of 20 OECD countries since 1980, we find that the growth rate of the age group of potential buyers (20-49 years old) actually experiences a strong cycle (due to migrations or to variations of the natural balance) that is the main driver of the ratio of housing investment to GDP. We use an instrument to identify causality and we find that, 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. Demographic changes are a better predictor of the housing investment rate than any other macroeconomic or financial variable we control for.

Although the long-term relationship between population and the housing stock is a well-known and obvious fact, this paper is the first to make the case that cyclical demographical change stands as an explanation of real housing business cycles.

Previous literature has ignored the short-term effect of demography on housing for two reasons:

First, the standard explanation of the volatility of housing investment is that housing is an asset. It follows that it is sufficient to look at the asset price (housing price) to explain the housing cycle, and that financial conditions are most likely to drive the cycle (Bénétrix, Eichengreen and O'Rourke (2012), Jordà, Schularick and Taylor (2015), Favara and Imbs (2015), Favara and Imbs (2015)). However, there are many reasons to believe that changes in housing prices do 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), Leamer (2007)).

Second, the literature on housing and demography only views the relationship between the two as a long-term relationship (Mankiw and Weil (1989), Lindh and Malmberg (2008)) Previous studies use the age group in level, a dependency ratio (the number of "young" people divided by the number of "old people"), or the share of the working age population in the total population. They are not able to capture the cyclical demographic effects because these variables experience a downward trend since the late 1980s.

Departing from the previous literature, this paper makes two main contributions:

First, based on a panel of 20 OECD countries since 1980, we show that the growth of population aged 20-49 is cyclical and 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 correlation 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 ask whether the observed correlation can be interpreted as a causal relationship between demography and housing construction. We propose a simple method to isolate and estimate the exogenous effect of the growth of population aged 20-49 on real residential investment. The population dynamic is potentially endogenous to housing investment and prices, especially if housing is indeed the business cycle (Leamer (2007)): people would immigrate to countries with a more dynamic housing sector and higher GDP growth. To circumvent this potential endogeneity bias, we use the growth rate of the population aged 0-29, twenty years before, to predict the current growth rate 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. By nature however, our instrument is 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 is low. Accordingly, we find a causal significant effect of the growth rate of population aged 20-49 only in the sub-sample with low migration, whereas a significant correlation is observed in both sub-samples. The difference between the two groups highlights that there is a potential sizeable endogeneity bias in the presence of high migration.

Overall, in the sub-sample where we are able to identify a causal effect with a strong instrument, we find that, 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.

In more intuitive terms, the conclusion of our study is that demographic changes contributed greatly to the housing construction cycle of the 2000s in OECD countries (both to the boom of the early 2000s and the bust of the late 2000s during the Great Recession) and that they would have been predictable using past demographic data, except in countries where migration was so high that it made such predictions impossible. Moreover, 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.

Beyond long term discussions, a better understanding of the housing cycle is of primary interest for economic modeling as well as for fiscal, monetary and macroprudential policies. Our empirical results suggest that demographic shock (by age) should be essential to any DSGE model aiming at reproducing housing investment cycles (relative to other components of GDP).

When discussing the causes of housing construction booms and busts, relative to GDP, policymakers should closely keep an eye on the demographic cycle. Financial variables and monetary policy (through inflation and nominal interest rates) may matter in some cases but, on average, they appear to play at most a minor role.

I. Introduction

Residential investment is one of 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 (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 do 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), Leamer (2007)). Then, what drives real investment in housing and why does it vary more than GDP? Beyond credit and asset prices, what about real factors?

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 correlation 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. These findings go against the common wisdom according to which - at the national level - population growth only affects the long-run housing stock but is not a main determinant of housing cycles. [1]

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. The population dynamics is potentially endogenous to housing investment and prices, especially if housing is indeed the business cycle (Leamer (2007)): people would 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). To circumvent this potential endogeneity bias, we use the growth rate of the population aged 0-29, twenty years before, to predict the current growth rate 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 is 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 is 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 sub-samples) 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 heteroscedasticity (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,

1. Studies that use a dependency ratio (the number of "young" people divided by the number of "old people"), or the share of the working age population in the total population are not able to capture such cyclical effects because these variables simply experience a downward trend since the late 1980s in OECD countries; see Section II.B.

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.

we find a causal significant effect of the growth rate of population aged 20-49 only in the sub-sample with low migration, whereas a significant correlation is observed in both sub-samples. The difference between the two groups highlights that there is a potential sizeable endogeneity bias in the presence of high migration.

In the sub-sample 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 in year t increases by 0.92pp (reduced form specification).[3] The instrumental variable estimator leads to a 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, the method used in this paper cannot distinguish between exogenous and endogenous effects of migrations. Despite such a limitation, our results have direct and important implications for macroeconomic policies and modeling, as well as for forecasts and debates on future GDP and housing growth, and on the impact of demography and migration on the housing market.

In more intuitive terms, the conclusion of our study is that demographic changes contributed greatly to the housing construction cycle of the 2000s in OECD countries (both the boom of the early 2000s and the bust of the late 2000s) and that they would have been predictable using past demographic data, except in countries where migration was so high that it made such predictions impossible.[4] 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). Conversely, the slowdown or decrease of the growth of the population aged 20-49 in OECD countries in the second half of the 2000s stands as

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. 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.

an explanation of the fall of the residential investment rate during the Great Recession. Furthermore, the current age structure is such that the United Nations predicts a 4% decrease between 2015 and 2045 of the numbers of people aged 20 - 49, in the 20 OECD countries included in our study.[5] Our findings suggest that these demographic changes are likely to have a negative impact on the residential investment rate, although it can potentially be mitigated by migrations.[6]

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 for the US as they observe a low correlation between interest rates and building permits at the national level.[7] 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 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)).[8] The direct implication of our paper for the DSGE literature is

5. See Figure B2 in Appendix B.

6. 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.

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

8. 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 hous-

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 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 II explores the several limits of the literature and highlights our contribution. Section III presents the benchmark estimations and main results. Section IV tackles the issue of endogeneity of the 20-49 age group's growth to housing conditions and explains our IV strategy. Section V provides a more thorough analysis of migrations, by age, overtime and by countries, and their correlations with the cycle of construction, in order to assess the potential size of the endogeneity bias.

II. Motivation and related literature

Residential investment is one of the most volatile components of GDP. In OECD countries, 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). [9] The standard explanation of such a peculiar pattern is that housing is an asset. As a consequence, studying the behavior of asset (house) prices is supposed to provide a comprehensive explanation of the volatility of housing flows (Poterba (1984),Topel and Rosen (1988)). In this framework, "the asset price of houses is a sufficient statistic for demand side forces, and the flow of new construction should therefore depend only upon the real house price" (Poterba (1984)). Thus, most of the studies on the housing

ing 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.

9. 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.

cycle look at the determinants of housing prices, which are viewed as a proxy for housing demand.

Accordingly, 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)). Since housing is an asset, loose financial and monetary conditions affect positively its demand and its price. Another strand of literature, starting with Mankiw and Weil (1989), explains housing prices mostly by demographic changes rather than financial factors. This literature has been 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)). [10] The literature which identifies housing demand with demographic factors either focuses on housing prices (Peek and Wilcox (1991), Takáts (2012), Eichholtz and Lindenthal (2014), Green and Hendershott (1996)) or aims at explaining directly both housing investment and housing prices (Mankiw and Weil (1989) , Lee et al. (2001), Lindh and Malmberg (2008)). Overall, the empirical debate on whether real housing prices is a sufficient statistic to study housing flows remains unsolved: neither the link between construction and prices, nor the link between prices and demography have been found to be robust. Yet, standard empirical macroeconomic studies of the housing market (DiPasquale and Wheaton (1994), Caldera and Johansson (2013)), based on the user-cost models, still use housing price as the dependent variable of the "demand equation" and housing investment as the dependent variable of the "supply equation". [11]

10. Agnello and Schuknecht (2011) and Bracke (2013) also control for the growth rate of the working age population whey they look for determinants of the turning points of housing price cycles.

11. As in Poterba (1984), the growth (or level) of housing investment and the housing investment rate are used indifferently, even though Poterba acknowledged that adjustment-cost theories of investment should favor the use of the rate of investment relative to the economy's total output.

In this paper, following the work of Mankiw and Weil (1989), we turn to a more empirical approach than most previous work on housing investment. Rather than relying on the user-cost model of Poterba (1984) which assumes that housing prices reflect perfectly the demand for housing, we will use demographic variables to identify changes in housing demand shocks and then to assess how they can explain the residential investment rate. Our motivation for such a specification is to explain why residential investment varies differently from the other components of GDP. In the following sections, we first explain why the housing price is not an appropriate dependent variable if we want to explain the peculiar dynamic of housing investment relative to GDP (Section II.A) and, second, we explain how the use of demographic variables as independent variables can identify changes in housing demand and why our strategy and explanatory variable differ from previous studies (Section II.B).

A. *Quantities and prices*

There are several empirical and theoretical reasons to believe that the asset price of houses may imperfectly reflect the cycle of the new flows that improve the housing stock.

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 of 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. Others, 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 because 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 actually differ in a significant way because of the peculiar structure of the housing market. 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 B5 in Appendix B). 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.

Fourth, housing prices series may provide important information on long-run evolution but, because of the imperfectness of prices data and of the differences in national sources overtime and across countries (Mack and Martínez-

García (2011), Knoll, Schularick and Steger (2014)), their short-run dynamic is unlikely to be comparable across countries. On the contrary, housing investment is calculated in national accounts and the definition is homogeneous across countries, at least from the 1980s.

B. *Identifying housing demand*

Mankiw and Weil (1989) use the US census to show that the purchase of a house is age specific, from which they conclude that the age profile strongly affects housing demand. [12] 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 III) whereas supply of new housing does not depend specifically on the growth of this age group. 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 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, and thus, may share similar properties (age, income etc.) with demanders. 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 even better in the case of housing investment than in the case of transactions or prices of existing homes (an argument that

12. 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.

was not developed by Mankiw and Weil (1989)). In addition, we can control for other changes in the population that could affect demand or supply of housing (growth rates of children and of older people, above 50). Finally, note that the dependency ratio (number of "young people" over the number of "old people") or the share of an age group in total population are not suited to identify precisely short-term variations in housing demand since these measures can be mostly driven by their denominators and thus only reflect long term trends in the population structure.

Previous literature considers population as a long-run determinant of housing demand (with the noticeable exception of the work of Muellbauer and Murphy (1997) on the UK housing prices). 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), and they looked at the correlation between this index and the level of residential investment. Departing from this literature, we look at the changes in population and we find - as an important stylized fact - that an age group growth can substantially vary and co-move with residential investment divided by GDP over time: Figure 1 shows - for some countries of our sample - that the growth of the population aged 20-49 and the residential investment rate experience a rather similar dynamic (see Figure B3 for the whole sample, and Figure B4 with the growth rate of the 30-49 in Appendix B). Such a cyclical variation is not observed with other demographic variables aiming to capture changes in population structures: the dependency ratio and the share of the population aged 20-49 in total population experience a downward trend since the 1980s because of an ageing population in OECD countries (cf. Figure B6 and Figure B7 in Appendix B). Surprisingly, the growth rate of age groups has not been used as an explanatory variable in previous papers on housing and demography. [13] Hence, they have been unable to capture the high volatility of housing demand. The high cyclicality of the 20-

13. Lindh and Malmberg (2008) divide the number of people of each age group by the total population and look at the correlation with the level of housing investment. Their paper is the closest to ours as it is the only one to study the link between demography on housing construction in a panel of OECD countries (over 1964 -1995) but they use different dependent and demographic explanatory variables, they only control for relative prices and they do not raise the issue of causation (cf discussion in the next section). Takáts (2012) uses the dependency ratio (the number of young people divided by the number of old people) to explain the changes in housing price, because he is interested in long term trends of the population structure (effect of ageing on asset prices) but not on the short-term dynamics of housing construction. Similarly,

49 age group's growth is a key fact highlighted in this paper. Cyclical changes in the growth rate of age groups can be due both to baby booms and busts or to changes in migration (as we will discuss further in Section IV). It suggests that a pure demand effect - based on fundamentals - might provide a sufficient explanation of the peculiar cycle of residential investment.

We are aware that the growth rate of the 20-49 age group is only a proxy (or an imprecise estimate) of the change in potential housing demand and is not a precise structural estimation of the potential housing demand. It is impossible to estimate exactly the potential housing demand for all the countries of our sample. Preferences, such as the age of the first house purchase or the age of leaving from the parents' house, are different between countries and change overtime. Moreover, those statistics are only available for some countries and for census period, and represent the actual (realized) average age of housing demand (and not the complete distribution) which may differ from the potential housing demand because of supply constraints. Results presented in the remaining of the paper are extremely similar if we use 1, 2 or 3 lags of the 20-49 growth rate, because growth rates of close age groups are very strongly correlated (this a cyclical variable).

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

Geerolf and Grjebine (2013) control for the dependency ratio to assess the impact of housing prices on investment and the current account.

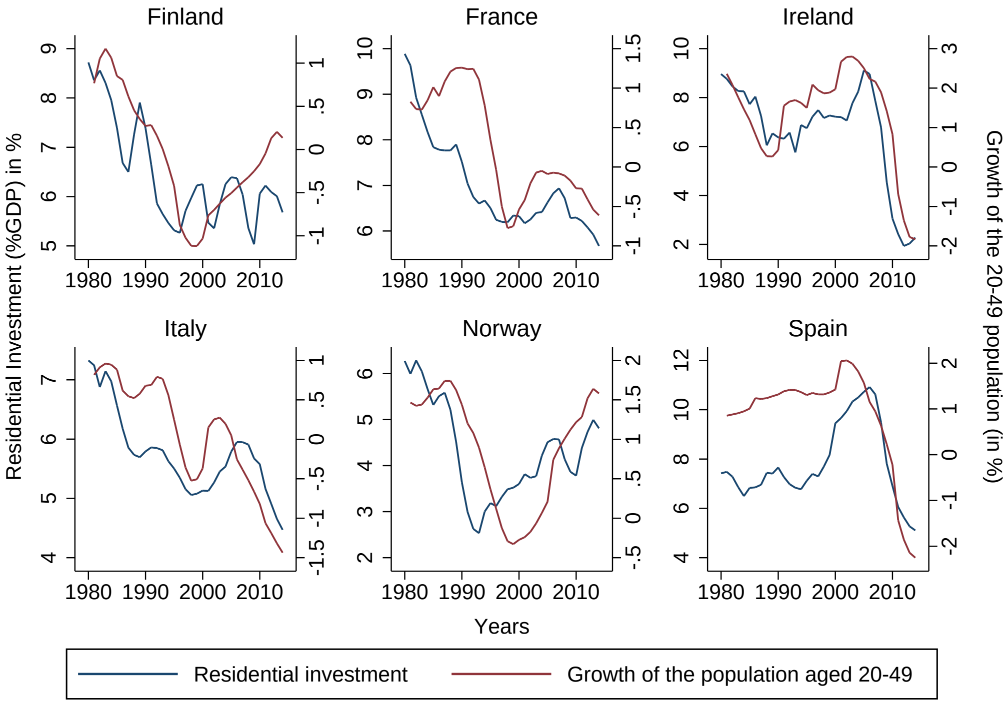


Figure 1. Residential Investment and the 20-49 population growth
For the whole sample, see Figure B3 in Appendix B

III. 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 the growth rate of several 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.[14] Expenditures in improvements and renovations are registered as investment if they increase the value of the building. Otherwise, they are registered as consumption (in construction).

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.[15] 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.

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).[16]

We scale 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. AMECO is the annual macroeconomic database of the European Commission's Directorate General for Economic and Financial Affairs (DG ECFIN). Sources of other macroeconomic variables used

14. 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).

15. 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.

16. 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.

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 represents 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 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 OECD database) to obtain the real volume of outstanding credit. Finally, the unemployment rate 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
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).

B. *Specification*

In order to study the correlation between the housing investment rate and changes in age groups, we estimate a balanced panel on 20 countries with annual data from 1980 to 2014. Despite being advanced economies, there are still large disparities in term of regulation on housing and mortgage markets between OECD countries. To account for this heterogeneity, we include country-fixed effects. We clustered our standard errors by country to correct for serial correlation and heteroscedasticity in the residuals. [17] To avoid the risk of spurious regression, we check the stationarity of our variables. [18]

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

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

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 [19], provided they are invariant over time, and $\xi_{c,t}$ is a disturbance term.

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

18. For the results of the tests, see Table C2 in Appendix C.

19. Given that the approach of this paper is to explain the business cycle of countries, it is natural to include country-fixed effects, but to avoid using time fixed-effects. Our objective is to control for as much macroeconomic variables as possible without overfitting the estimation with year dummies that would naturally capture common shocks and explain most time variation. Implications for macroeconomic modelling would be null if most of the business cycle is explained by time-fixed effects. The risk of overfitting would be especially damaging in our case because there is a common cycle in the growth rate of working age population in OECD countries (as shown on Figure B1 in Appendix B) . Only Ireland and Spain in the early 2000s offer exception to this common pattern. Year dummies may capture this common cycle of population and then housing demand. However, we have checked whether our main results are affected by the addition of time-fixed effects in our estimation. As seen on Table C5, the 20-49 age group is still significant but other age groups become significant (especially the youngest age group) because of the collinearity induced by the time fixed effects.

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, 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 remain stable. 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 also find that in our 20-country 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 2, 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 2 and discussed below (Section III.C), consumption and non-residential investment follow different patterns.

2. The choice of age groups

The choice of the number of age groups entails a trade-off. With 5-year age groups, the age group's coefficients will not be precisely estimated since age groups are rather correlated (people flow from one age group to the other as they aged). With only few age groups, there is a risk of an arbitrary choice. [20] To choose a limited number of age groups as independent variables on a robust empirical basis, we first estimate the relation for 6 age groups (cf. Table 3) and then aggregate the age groups that have a similar impact on residential investment after checking that multicollinearity is not an issue. Among our six age groups, only the 20-29 and 30-49 have a positive significant effect. Hence, based on these results, we finally divide the population in three age groups (cf. Table 4): the 0-19 years old, the 20-49 years old, and the 50 years old and above. As explained in Section II.B, our goal is to focus on an age-group of people whose growth affects housing demand differently from the supply of new housing. Then the other age groups (0-19 and 50+) are still used as controls in the regressions as they can also influence the demand or supply of housing.

Finally, we check that it is really the age distribution that impacts the ratio of the residential investment and not the total population growth. Indeed, as it can be seen on Table C4, the growth of the total population is not significantely correlated with the residential investment ratio.

3. Controls

Theoretically, the demand of housing should increase with income growth. The personal disposable income represents the income resources "at hand" for

20. In the literature, the common solution is to use one variable aggregating all the information of the distribution such as dependency ratios (Takáts (2012)) but it suffers from the problem highlighted in Section II.B. Another option is to impose a polynomial restriction on the coefficients (Fair and Dominguez (1991) or Higgins (1998), or in the case of Mankiw and Weil (1989)) or to impose 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.

Table 2 – 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.

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)).[21] 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*

Table 4 show our main results when using 3 age groups as explanatory variables. As expected, the growth rate of the population aged 20-49 is positively correlated with the residential investment rate. In Table 4, we take column (5) as our benchmark, since we do not have sufficient observations for credit in the case of Luxembourg. Column (6) includes credit growth as a control but excludes Luxembourg. In column (5), an increase of one percentage point in

21. 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 3 – 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 4 – 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.

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.

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 has a negligible impact on residential investment (when household's real credit growth increases by one percent, residential investment increases by 0.07 pp cf. Table C6 in Appendix C). As expected, the real housing price growth has a positive impact on residential investment but it is very small too: when prices increase by 1%, the residential investment rate only increases by 0.04 percentage points. [22] The variation of unemployment has the expected negative sign, 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 residential investment differently from 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. The bias of their coefficients is not transmitted to the coefficients of demographics variables, because they remain stable throughout the specifications, as shown in the different columns of Table 4. In particular, adding the growth of housing prices only lowers the coefficient of the 20-49 age group from 1.02 to 0.98 and increases the R^2 from 0.27 to 0.3. Overall, the

22. It means that a 13.3 pp increase of the growth of real housing prices would be necessary to obtain an increase of the residential investment rate of 0.5 pp (whereas the actual average growth rate is 1.46 % and standard deviation is 6.41). The same effect is obtained by a 0.55 pp increase of the growth rate of population aged 20-49, which the average value of this variable in our sample.

controls are weakly correlated with the growth of the age group 20-49 [23] and we do not find the presence of multicollinearity in the regression. [24]

D. *Discussion and robustness*

Is the age profile effect on housing investment also observed on other components of GDP? Table 2 shows 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 (which is confirmed by Figure B8 in Appendix B). 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, when we replace the dependent variable by consumption or 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 C7 in Appendix C): the coefficients of the growth of the age group 20-49 are negative and insignificant.

One can argue that investment in housing could be substitutable with other forms of savings (even though, at the aggregate level, total saving obviously equals total investment). In presence of such a substitution, the recent decrease of residential investment over GDP could be driven by an increase in aggregate savings. It could be the case, for example, if there is an increase of precautionary savings (more liquid than housing) by the 20-49 years old, who expect lower retirement pensions in the future. Indeed, Giavazzi and McMahon (2008) show that an increase in uncertainty about the future path of income will significantly increase precautionary savings. Financial liberalization - that greatly

23. 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).

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

increases in OECD countries since 1980 - could favor such a substitution effect if it provides more financial alternatives (such as life insurance), as underlined by Schmidt-Hebbel and Servén (2002), to housing investment for precautionary savings. We check whether this mechanism holds in our sample. In Table C8 in Appendix C, we observe that saving as a share of GDP is positively correlated with residential investment ratio, which refutes a potential substitutability between the two variables and suggests that we might simply capture the identity between saving and investment. Moreover, the coefficient on the 20-49 years old population's growth remains high and significant.

In Section III.C, we have seen that the inclusion of house prices as a control variable does not add explanatory power in the estimation. What about the potential correlation between demography and house prices? As we argue in Section II.B, since most of our price data is about existing dwellings, we could catch a supply effect when regressing the age distribution on prices. Indeed, as it can be observed on Table C9, we find a significant impact of the 50 years and above age group. This effect could be partly due to the fact that we catch a supply effect related to the selling of their existing dwellings by the middle-aged. In any case, we cannot find a robust link between housing demand and house prices, contrary to what we found between housing demand and the residential investment rate.

In the previous estimations and argumentation, we have not taken into account the household size. Household size is however an obvious and well 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. Neither macroeconomic cross-country studies, nor national case studies, have 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. In what extent is it likely to bias our results?

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 growth rate 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 B10 in Appendix B. Overall, Figure B10 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 OECD 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. Despite a clear long-term downward trend explained by sociological factors, household size could however experience short-term variations (partly unobserved in official data) due to economic or housing conditions. For example, young adults would live in their parents' house for a longer period if the housing market is tight or if their unemployment rate is high. If this is the case, our controls (housing prices, unemployment rate and disposable income) would capture such effects. Thus, we do not find evidence that the inability 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.

IV. 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 if residents emigrate when a country experiences low growth (which can be due or correlated to a housing crisis), then it is impossible 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 and more job opportunities (correlated with a booming economy and housing sector), 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 all the more important for our estimations since a major share of migrants are likely to be between 20 and 50 years old. [25] 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 macroeconomic paper discussing or attempting to cope with this issue. Furthermore, since there is no published statistics on the age of the flow of migrants (see Section V), 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 seen 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 constant mortality for each age group over time, the growth of the population aged 20-49 today will be approximately equal to the growth of the population aged 0-29, 20 years ago. And the latter variable affects the residential investment rate in t only through

25. 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 former variable. 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 both to discuss whether the endogeneity problem is potentially sizeable (that is if migrations contribute significantly to the growth rate of the age group) and to provide a causal examination of the impact of demographic changes on housing demand.

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 percent 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 percent of the population aged between 0 and 29 years old at time $t-20$ in country c. Equation (1) still includes the controls ($X_{c,t}$) of the previous estimations 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. [26] The Spanish and Irish cases in the early 2000s are

26. Note that we use the term "population flows" or "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 most striking (Figure 2a): in these countries, the annual 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 effect" in Europe that would have caused a high increase in net migrations in the Schengen area countries only, beginning in the 1990s. Indeed, for example, the 1998 population projections of the United Nations (based on past migration data) predicted a decrease of the people aged 20-49 in the UK, Norway and Spain, between 2000 and 2007, whereas their number actually grew substantially (cf. Figure B9 in Appendix B).

As a consequence, there is no evidence that a common migration shock would have affected some or all countries of the sample under study between 1980 and 2014. There is also no evidence that countries systematically differ in their level of migrations over the whole sample. As seen on Figure 2a, there were migration booms, but in some countries only, and only concentrated in a small part of the sample. Thus, the level of migrations is more likely to be related to country-specific characteristics which varied overtime. Such time-varying country-specific characteristics of migration suggest further potential evidence of an endogenous link between migration and the business (and housing) cycle.

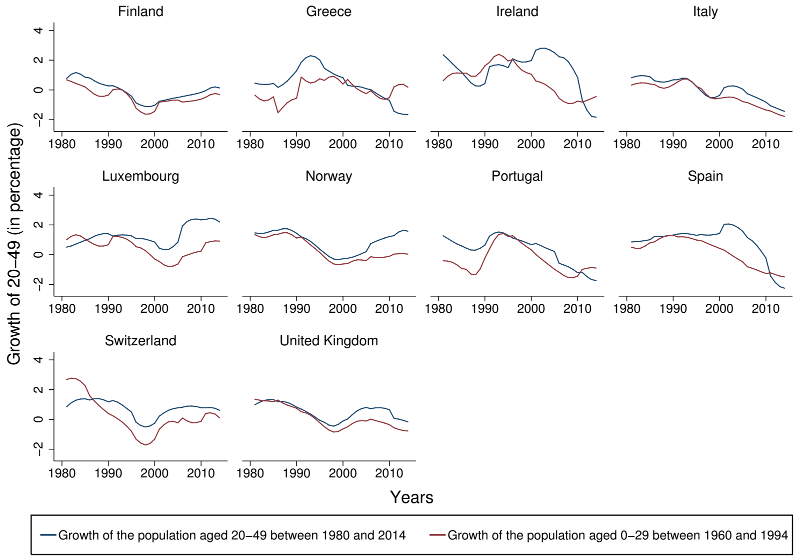
The previous discussion highlights an important characteristic of our instrument. It will be a strong instrument except for countries that had experienced substantial net migration. The instrument is expected to be weak in countries where, on average, the actual growth of the 20-49 significantly differs 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 sub-sample, 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 sub-sample is called "Low migration group" (Figure 2b) whereas the other sub-sample is called "High migration group" (Figure 2a). The rationale for simply dividing the whole sample in two 10-country sub-samples is twofold. First, as seen on Figure 2a, there are 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 sub-samples. 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*

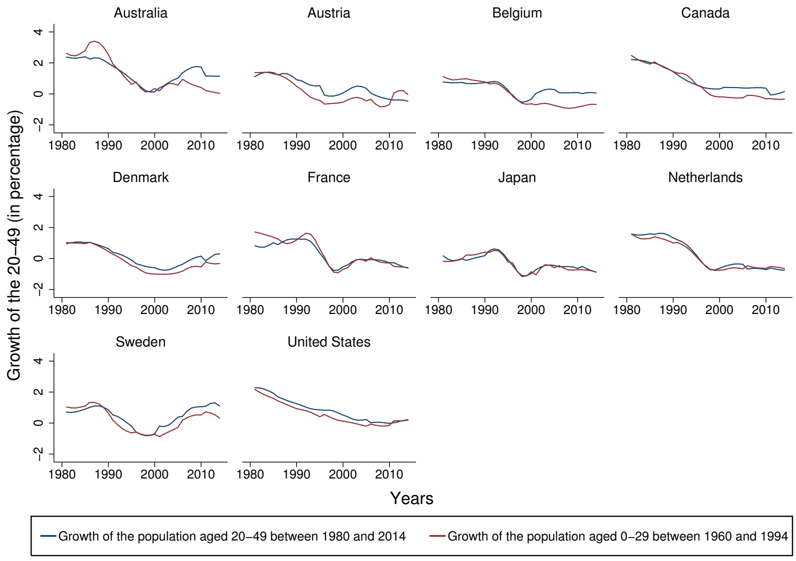
We first report the results of the reduced form estimation (Table 5).[27] The reduced form is directly interpretable and interesting in itself since it measures the causal impact of the predicted growth of population aged 20-49 under the assumptions of zero net migration and a constant mortality rate by age group. In the sample of countries that experienced low population flows, the predicted growth of the 20-49 years old without migration (proxied by the growth rate of population aged 0-29 in t-20) does well in explaining the current variation of the residential investment. Indeed, we find that an increase of 1 pp of the growth rate of the population aged 0-29, twenty years before, would increase the residential investment rate in t by 0.92 pp. However, in the sub-sample with high-migration, 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 sub-sample, 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, 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 find a lower impact, although still significant.

Table 6 displays the results of the IV regression. The Montiel-Pflueger test assesses that, as expected, the instrument is strong in the "Low migration sample" but weak in the "High migration sample". It also points to a strong

27. 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.

instrument for the whole sample.[28] 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 is less than one (cf. Table C10 in Appendix C). 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 significant but lower (0.65) than with the "Low migration" sample. As with the reduced form, the impact is not significant 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 residential investment rate than any other macroeconomic variable.

C. *A further look at subsample correlations*

It is interesting to run our initial regressions of Section IV.B , without instrumentation, on each sub-sample in order to assess the size of the potential endogeneity bias (Table 7), as we expect a higher correlation in presence of endogeneity. The size of the bias actually looks rather small, yet non-negligible: the coefficient of the growth of population aged 20-49 is 0.93 in the low migration group and 1.10 in the high migration group (the difference between the two coefficients is significant).[29]

Furthermore, Table 7 shows that, 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 residential investment rate in countries that

28. 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)).

29. Due to the difficulty of doing a proper Wald test between different panel regressions as underlined by Blackwell III (2005), we do a Z-test to test for the significativity of the difference between two coefficients : $Z = \frac{b_1 - b_2}{\sqrt{SEb_1^2 + SEb_2^2}}$ where $b_1$and b_2 are different coefficients and SE represent their standard errors. In the following of the paper, we implicity refer to this test when we talk about significant difference between coefficients of different regressions.

did not experience high migration. In those countries, it would have been possible to forecast the actual growth of the 20-49, and, accordingly, to foreseen quite accurately the path of the residential investment rate. On the contrary, 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. In that respect, 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.

The next section aims to decompose more precisely population growth between migration and natural balances in order to both bring more robustness to our previous estimations and assess further the size of the potential endogeneity bias.

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.

V. Migration by age

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

Because of migration, there is a large difference - in some countries of our sample - between $\Delta Pop0\text{-}29_{t-20}$ and $\Delta Pop20\text{-}49_t$. We now want to investigate whether, in the high migration group, residential investment is indeed correlated with migration flows of people aged 20-49, as well as to assess the importance of the endogeneity bias from the size of the correlation. We are still unable to identify a causal effect of migrations flows because there is no good instrumental variable to predict migration flows at the national level, as underlined by McKenzie and Sasin (2007). National and international statistical institutes do not publish data on migration flows by age groups. They usually estimate 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. The estimation of migration flows in this section aims to calculate the magnitude of net migration for each year and for each age group. Moreover, by reconstructing the migration by age, we take into account the potentially time-varying mortality of age groups. Hence, our new estimation of annual net migration by age is different from the net migration over twenty years that could be inferred from the reduced form estimation of the previous section. [30]

30.

$$\begin{aligned} Pop^x_t - Pop^x_{t-1} &= Netmigr^x_{t-1,t} + Naturalbal^x_{t-1,t} \\ Pop^{x-20}_{t-20} - Pop^{x-20}_{t-21} &= Netmigr^{x-20}_{t-21,t-20} + Naturalbal^{x-20}_{t-21,t-20} \end{aligned}$$

where the first equation represents the decomposition of the population growth between net migration and natural balance of the people aged x in t, and the second equation represents the same thing for people aged $x-20$ in $t-20$. Substracting the two equations, the natural balances cancel out since they concern people born in the same year, we obtain :

$$\triangle Pop^x_t - \triangle Pop^{x-20}_{t-20} = Netmigr^x_{t-1,t} - Netmigr^{x-20}_{t-21,t-20}$$

if the growths of the age groups twenty years apart are equal, it means that there were no new net migrations since the initial period, or in other words, the size of the net migration during 20 years, of the people aged less than 20 years old, did not significantly increase the growth of the 20-49 population today.

To estimate net migration by age, we use the forward method of the cohort component population model, 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 obtain population data and survival rates for each age annually (cf. data sources in Appendix A), 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, subtracting 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 on annual data, this assumption is not crucial for the results: it only excludes migrants who died during the year of their migration.

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 errors in the census counts or to definition or legal changes from one census period to the next. There are also differences in measurement, some countries have a resident registry such as Austria, Denmark, Finland, Norway, Netherlands, Italy, Luxembourg and Japan whereas others do not force their residents to register. Moreover, when estimating population for each age, statistical institutes themselves make assumptions about migrations that we actually recover by applying our method. Due to data availability, we can only apply our method from 1986.

B. *Specification and results*

We divide our estimation of net migration for population aged 20-49 by the total population aged between 20 and 49 years old, in order to have coefficients comparable to the 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:

$$
\begin{aligned}
GFCFdwellings/GDP_{c,t} &= \alpha_0 + \alpha_1 \frac{NetMigr20\text{-}49_{c,t}}{Pop20\text{-}49_{c,t-1}} + \alpha_2 \Delta Pop0\text{-}29_{c,t-20} \\
&+ \alpha_3 \Delta Pop0\text{-}19_{c,t} + \alpha_4 \Delta Pop50+_{c,t} + \alpha_5 X_{c,t} + \gamma_c + \xi_{c,t}
\end{aligned}
$$

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, and $\Delta Pop0\text{-}29_{c,t-20}$ is the growth of the population aged 0-29 years old between 1960 and 1994. $\Delta Pop0\text{-}19_{c,t}$ 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 controls and the country fixed-effects.

Note that we control by $\Delta Pop0\text{-}29_{c,t-20}$ rather than by a measure of the natural balance of population aged 20-49, defined as $NatBal20\text{-}49_{c,t} = Pop20\text{-}49_{c,t}$-$NetMig20\text{-}49_{c,t}$, and scaled by $Pop20\text{-}49_{c,t}$. The rationale for doing so is to keep an exogenous proxy of the natural balance as an explanatory variable. Indeed, if someone immigrates before 20 years old, it will be counted as part of the natural balance of the 20-49 years old in the subsequent years. Consequently, $NatBal20\text{-}49_{c,t}$ would include partly the evolution of recent net migration and could not be considered as truly exogenous.

The results in Table 8 confirm the results of Section IV. In the countries that did not experience substantial net migration flows, residential investment cycle is explained by $\Delta Pop0\text{-}29_{c,t-20}$ and not by migration flows (the coefficient of 0.90 is not statistically different from the coefficient of 0.92 of the reduced form estimate, in the case of low migration countries). In the high migration countries, an increase of one percent of the contribution of net migration to the growth of the 20-49 population increases residential investment by 0.85 pp ; the net migration dynamics in high migration countries have a similar

Table 8 – Residential investment, migration and natural balances

	(1) All sample	(2) Low population flow	(3) High population flow
Constant	6.27***	4.91***	7.49***
	(0.77)	(0.95)	(0.29)
Migration balance			
20-49	0.42	−0.64	0.85***
	(0.28)	(0.39)	(0.19)
Natural balance			
0-29 (1960-1994)	0.29	0.90*	−0.25
	(0.33)	(0.45)	(0.30)
Population growth			
0 - 19	0.12	0.45	−0.03
	(0.21)	(0.26)	(0.34)
50 plus	−0.38	0.43	−1.19***
	(0.55)	(0.57)	(0.20)
Controls			
Real House price growth	0.04***	0.05**	0.02*
	(0.01)	(0.02)	(0.01)
Real Disposable income growth	0.03	0.07	0.03
	(0.04)	(0.04)	(0.06)
Real long term interest rate	−0.04***	−0.02	−0.04***
	(0.01)	(0.04)	(0.01)
ΔUnemployment	−0.13	0.10	−0.24
	(0.14)	(0.12)	(0.16)
Observations	560	280	280
No. of countries	20	10	10
R^2 within	0.25	0.28	0.37

Note: Panel linear regression of the percentage share of GFCF dwelling on GDP, both in volumes, on the migration balance of the 20-49 years old, the growth of the 0-29 years old between 1960 and 1994, 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.

impact on residential investment as $\Delta Pop0\text{-}29_{c,t-20}$ in low migration countries (the 0.92 coefficient in the reduced form regression for low migration countries is not significatively different from the 0.85 coefficient on net migration for the high migration countries in this regression). As expected, if we use $\frac{NatBal20\text{-}49_{c,t}}{Pop20\text{-}49_{c,t-1}}$ instead of $\Delta Pop0\text{-}29_{c,t-20}$ in the regression, the effect of the former is higher because this variable incorporates recent net migration flows and is correlated with net migration dynamics (cf. Table C11 in Appendix C). To put in a nutshell, even if we cannot assert the causality between migration and the residential investment rate, we find that the correlation between net migration by age and residential investment is of similar size as the effect of exogenous variations of the age group 20-49. [31] These results suggest that the endogeneity bias is in fact small or null, although it is potentially important given the large contribution of migration to the growth of population 20-49 in some countries.

VI. Conclusion

Common wisdom suggests that population growth is an important determinant of the housing stock in the long-run. In this paper, we argue that demographic changes are also crucial to understand the cyclical dynamic of housing investment and especially why it differs from the dynamic of other components of GDP. The cyclicality of the residential investment rate is well explained by the growth rate of the population aged 20-49. Indeed, the latter is more volatile than commonly thought, because of baby booms and busts as well as changes in migration. The growth rate of the 20-49 age group affects the demand for housing differently from the other components of GDP. An increase of one percentage point of the 20-49 years old growth rate implies an increase of 1.3pp of the ratio of residential investment over GDP. This finding is extremely robust to many controls and specifications, and we have shown that it can be identified as a causal effect. The correlation is also observed for migrations (of people aged 20-49), but the causality cannot be assessed in this case. Our decomposition between migration and natural balance by age shows that the endogeneity bias is potentially important because of the large contri-

31. The net migration in $t-2$ is the lag that has the highest coefficient, whereas the highest coefficient in the natural balance case is the contemporary one. This reinforce the evidences in favor of a small endogeneity bias. Net migration could create an additional housing demand, and housing investment adjusts with a lag.

bution of migration to the growth of the age group 20-49, but comparisons of the estimates suggest however that the size of the bias is in fact not large.

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 ratio of residential investment to GDP growth is actually low (6% 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 and on employment, 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 economic modeling as well as for fiscal, monetary and macroprudential policies. Our empirical results suggest that demographic shock (by age) should be essential to any model aiming at reproducing housing construction cycles. When discussing the causes of housing construction booms and busts, relative to GDP, policymakers should closely keep an eye on the demographic cycle. Financial variables and monetary policy (through inflation and nominal interest rates) may matter in some cases but, on average, they appear to play at most a minor role.

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.

Blackwell III, J Lloyd. 2005. "Estimation and testing of fixed-effect panel-data systems." *Stata Journal*, 5(2): 202–207.

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."

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.

DiPasquale, Denise, and William C. Wheaton. 1994. "Housing Market Dynamics and the Future of Housing Prices." *Journal of Urban Economics*, 35(1): 1–27.

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.

Geerolf, François, and Thomas Grjebine. 2013. "House Prices Drive Current Accounts: Evidence From Property Tax Variations." CEPII research center Working Papers 2013-18.

Giavazzi, Francesco, and Michael McMahon. 2008. "Policy Uncertainty and Precautionary Savings." C.E.P.R. Discussion Papers CEPR Discussion Papers 6766.

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.

Harris, Richard DF, and Elias Tzavalis. 1999. "Inference for unit roots in dynamic panels where the time dimension is fixed." *Journal of econometrics*, 91(2): 201–226.

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.

Im, Kyung So, M Hashem Pesaran, and Yongcheol Shin. 2003. "Testing for unit roots in heterogeneous panels." *Journal of econometrics*, 115(1): 53–74.

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.

Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu. 2002. "Unit root tests in panel data: asymptotic and finite-sample properties." *Journal of econometrics*, 108(1): 1–24.

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.

McKenzie, David, and Marcin J. Sasin. 2007. "Migration, remittances, poverty, and human capital : conceptual and empirical challenges." The World Bank Policy Research Working Paper Series 4272.

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.

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

Schmidt-Hebbel, Klaus, and Luis Servén. 2002. "Financial liberalization, saving, and growth." *Macroeconomic Stability, Financial Markets, and Economic Development.*

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.

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 Quarterly) 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 Quarterly 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 hyperinflation). 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 variables 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 labor 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.

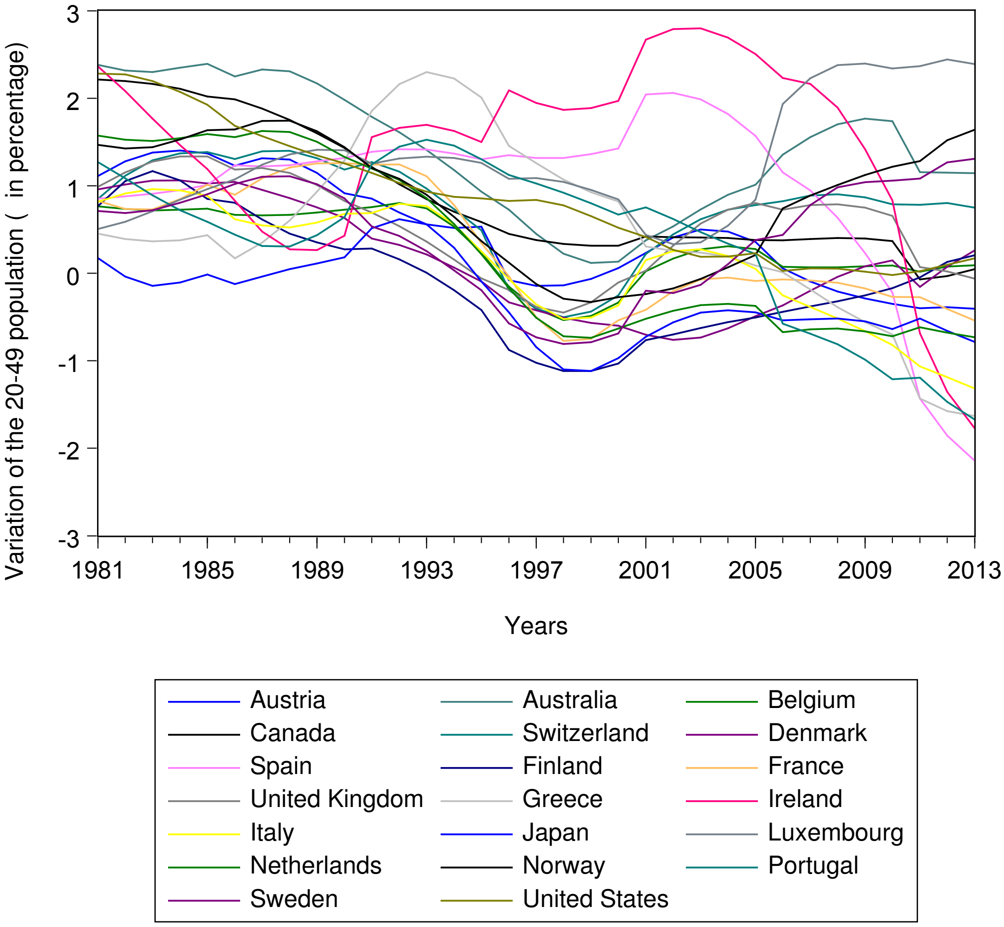


Figure B1. Growth of the population aged between 20-49 years old

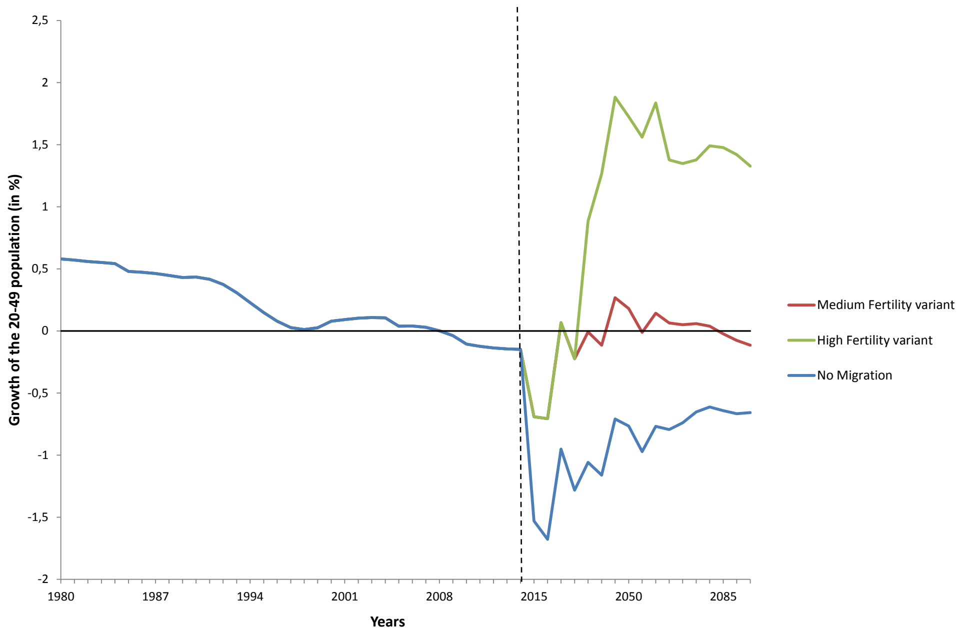


FIGURE B2. FORECAST OF THE 20-49 FROM 2015 TO 2100

Source: UN World Population Prospect

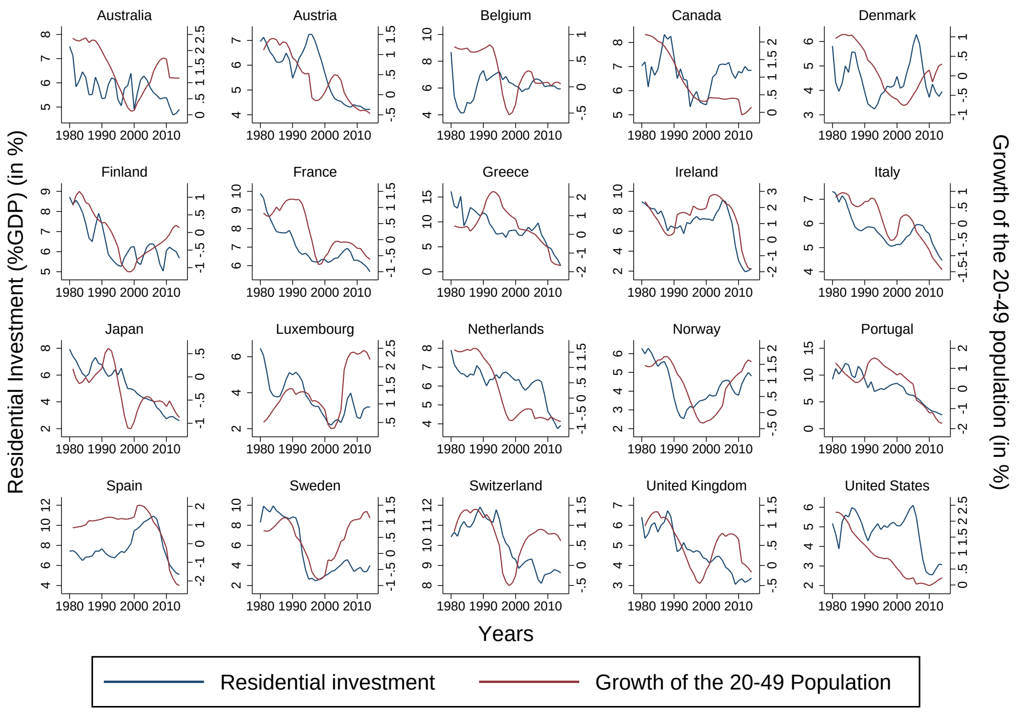


FIGURE B3. RESIDENTIAL INVESTMENT (%GDP) AND GROWTH OF THE 20-49

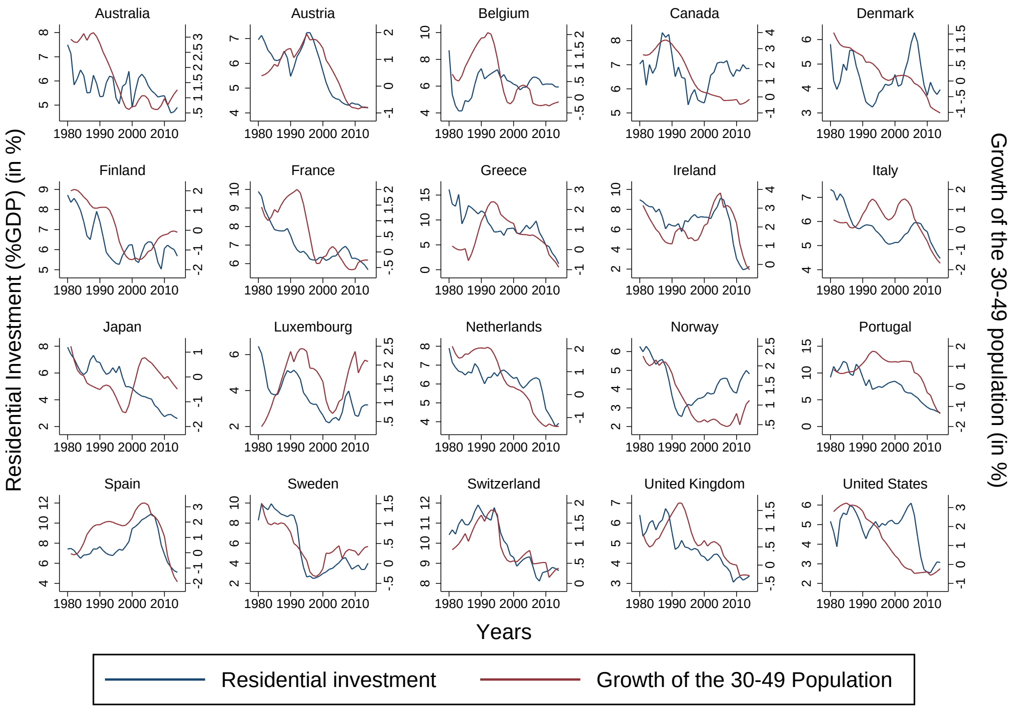


FIGURE B4. RESIDENTIAL INVESTMENT (%GDP) AND GROWTH OF THE 30-49

FIGURE B5. GROWTH OF RESIDENTIAL INVESTMENT VS HOUSING PRICES

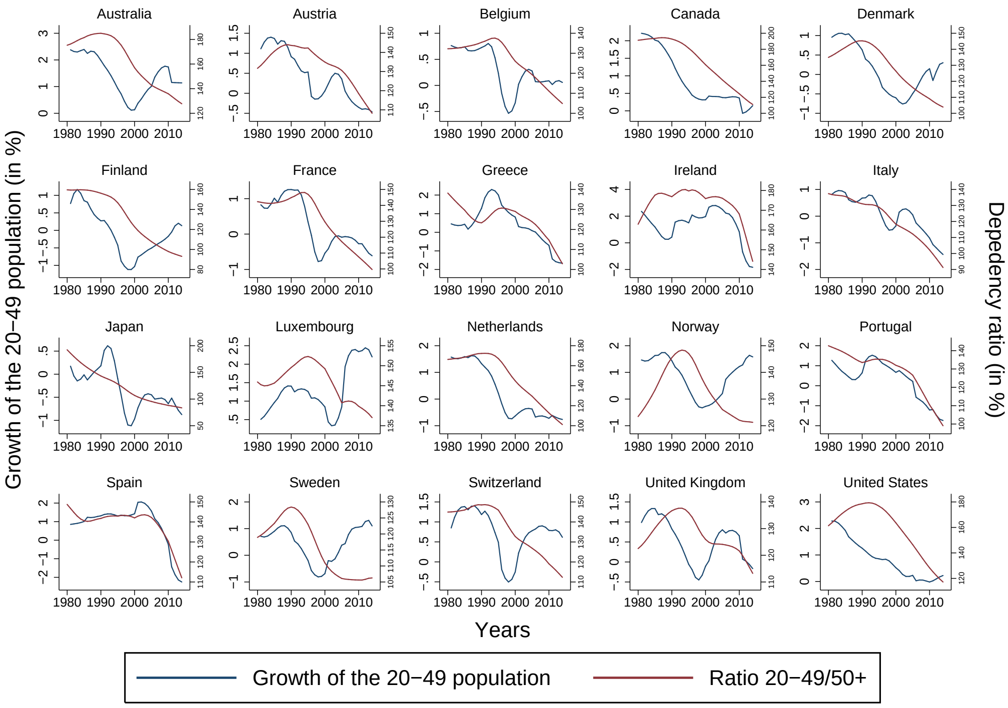


FIGURE B6. GROWTH OF THE 20-49 VS DEPENDENCY RATIO

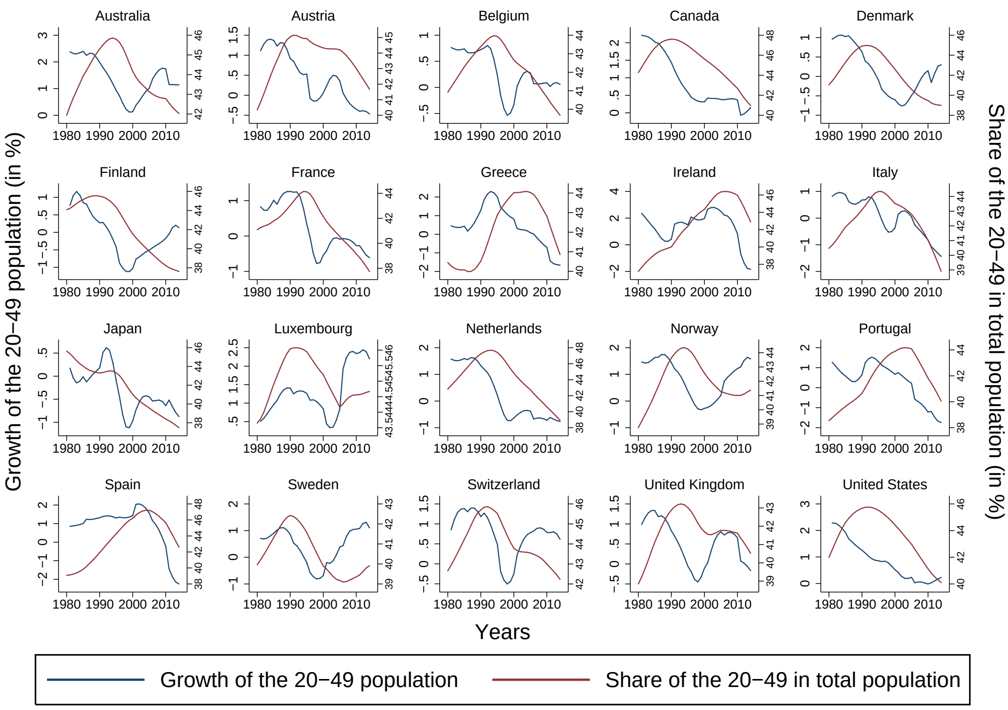


FIGURE B7. GROWTH OF THE 20-49 VS SHARE OF THE 20-49 IN THE POPULATION

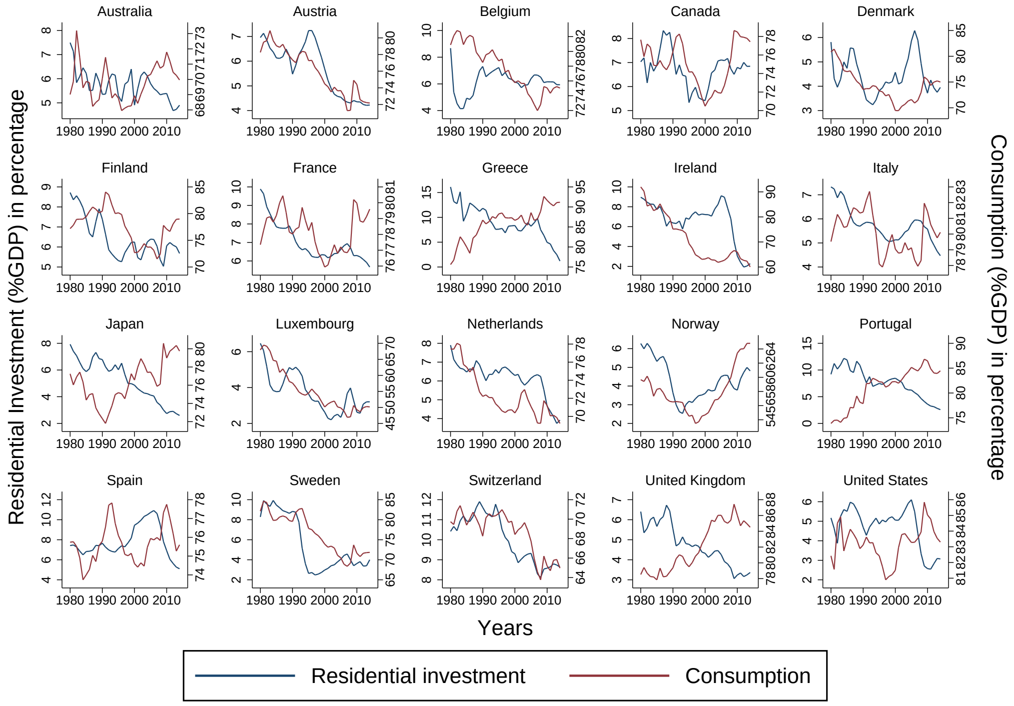


FIGURE B8. RESIDENTIAL INVESTMENT AND CONSUMPTION (%GDP)

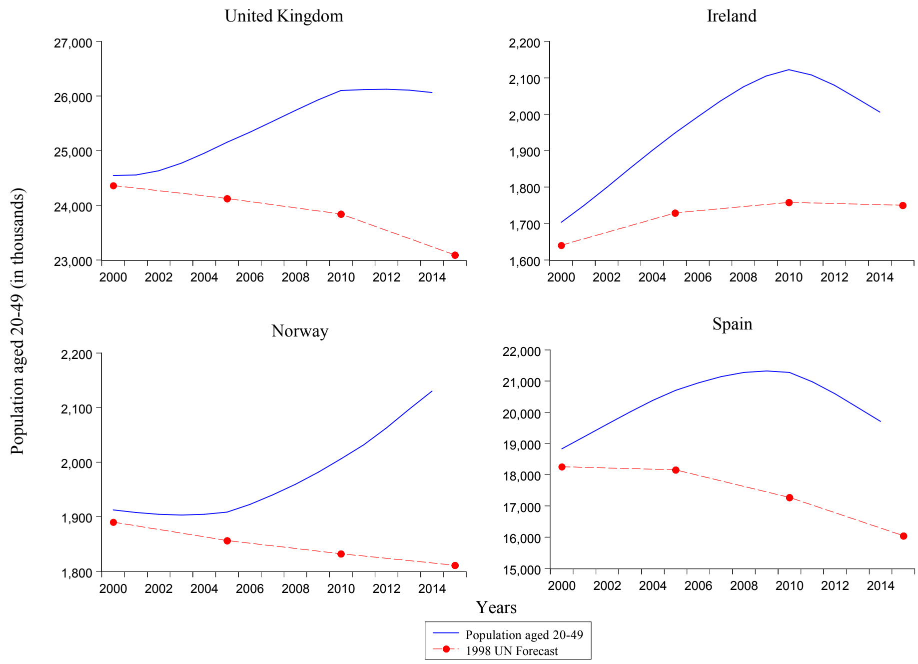


FIGURE B9. 20-49 POPULATION AND ITS FORECAST IN 1998

Source: UN Demographic Yearbook 1998

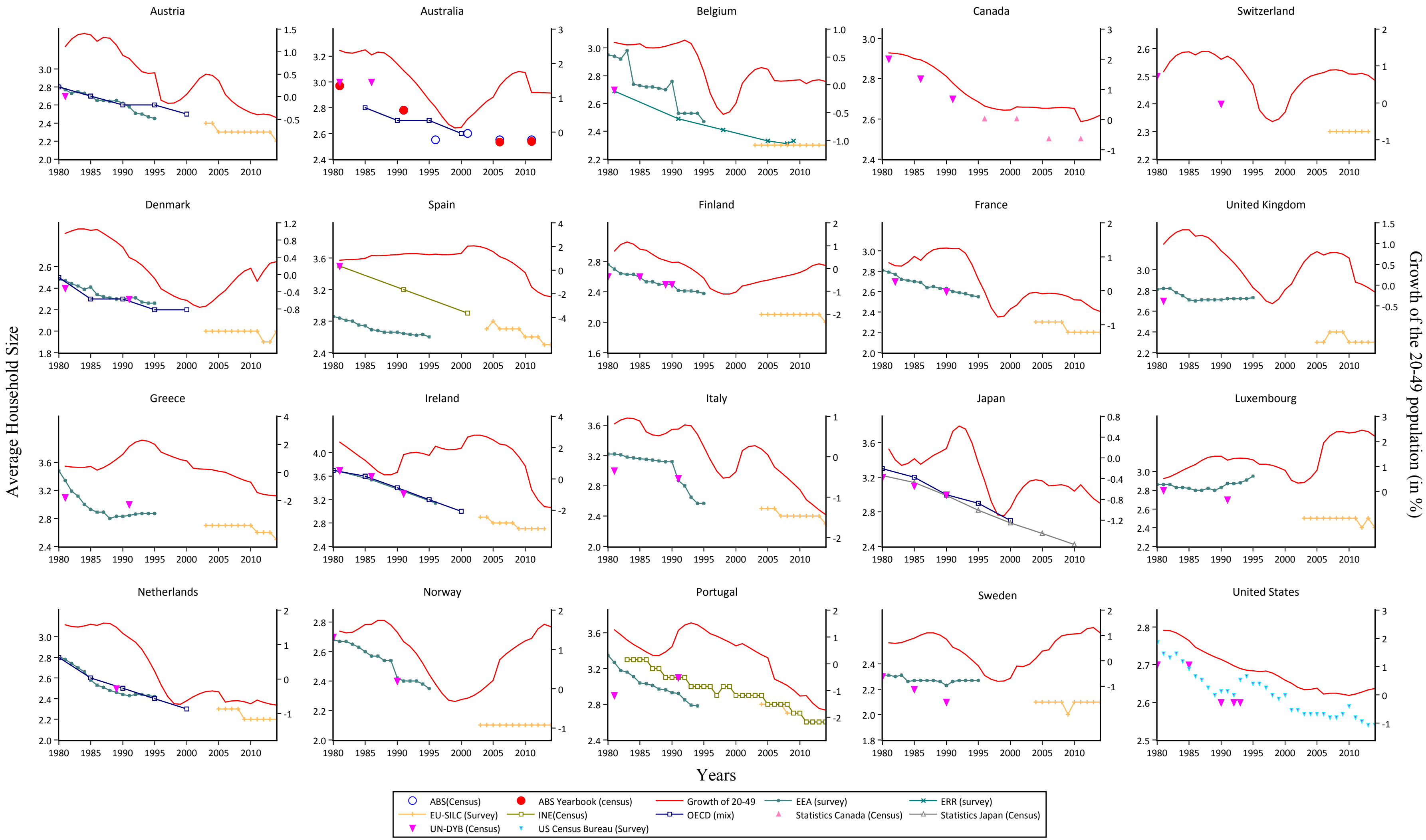


Figure B10. Average Household Size and Growth of the 20-49 Population

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 – Panel Unit Root test

	P-values Im, Pesaran and Shin (2003)
Main variables	
Residential investment	1,00
Residential Investment (%GDP)	0,06
ΔPopulation 0-19	0,00
ΔPopulation 20-49	0,05
ΔPopulation 50 plus	0,51
ΔPopulation 0-29 (1960-1994)	0,00
ΔNet migration	0,01
Controls	
ΔReal disposable Income	0,00
ΔReal House Price	0,00
Real Long Term Interest Rate	0,00
ΔUnemployment	0,00
ΔReal Credit	0,00

Note: H_0: all panels contain unit roots.

The Im, Pesaran and Shin (2003) test (IPS in the following) is based on the augmented Dickey-Fuller regression and allows for serially correlated and heteroskedastic errors (which is the case in our study as seen in Table C1). The starting point for the test is a set of Dickey–Fuller regressions of the form:

$$\triangle y_{it} = \phi_i y_{i,t-1} + \gamma_i + \sum_{j=1}^{p} \triangle y_{i,t-j} + \epsilon_{it}$$

where $i = 1, ..., N$ indexes panels and $t = 1, ..., N$ indexes time. y_{it} is the variable that being tested, and ϵ_{it} is a stationary error term. γ_i represents panel-specific means (fixed effects). The null hypothesis of the presence of unit roots is to test $H_0 : \phi_i = 0$ for all i versus the alternative $H_a : \phi_i < 0$. Most panel unit-root tests such as the Levin, Lin and Chu (2002), the Harris and Tzavalis (1999) and the Breitung (2000), make the simplifying assumption that all panels share the same autoregressive parameter so that $\phi_i = \phi$ for all i, whereas the IPS test allow the autoregressive parameter to be panel specific. This regression is applied to each panel separately and then Im, Pesaran and Shin (2003) average the resulting t statistics. The number of lags p is chosen such that it minimizes the AIC for the Dickey Fuller regression. For each time period, we also compute the mean of the series across panels and substract this mean from the series. Levin, Lin and Chu (2002) suggest this approach to mitigate the impact of cross-sectional depedence (which is present in our sample as seen in Table C3).

Table C3 – 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.

Table C4 – Total population growth vs age group growth

	Residential investment(% GDP)	
	(1) Total population	(2) With age groups
Constant	5.61***	5.95***
	(0.28)	(0.59)
Population growth		
Total Population	0.79	
	(0.50)	
0 - 19		0.21
		(0.18)
20 - 49		0.97***
		(0.20)
50 plus		−0.20
		(0.35)
Controls		
Real Disposable income growth	0.01	−0.02
	(0.04)	(0.04)
Real House price growth	0.04**	0.04**
	(0.02)	(0.01)
Real long term interest rate	−0.05***	−0.03***
	(0.01)	(0.01)
ΔUnemployment	−0.03	−0.09
	(0.09)	(0.07)
Observations	680	680
No. of countries	20	20
R^2 within	0.11	0.31

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 or the total population growth, 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 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.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 – Country Fixed Effects vs Time Fixed Effects

	Residential investment(% GDP)	
	(1)	(2)
Constant	5.95***	7.19***
	(0.59)	(0.51)
Population growth		
0 - 19	0.21	0.50***
	(0.18)	(0.16)
20 - 49	0.97***	0.57***
	(0.20)	(0.18)
50 plus	−0.20	0.40*
	(0.35)	(0.23)
Controls		
Real Disposable income growth	−0.02	−0.02
	(0.04)	(0.03)
Real House price growth	0.04**	0.02
	(0.01)	(0.02)
Real long term interest rate	−0.03***	−0.03***
	(0.01)	(0.01)
ΔUnemployment	−0.09	−0.13*
	(0.07)	(0.07)
Observations	680	680
No. of countries	20	20
R^2 within	0.31	0.52
Time fixed effects	No	Yes

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 only have country fixed effects. In column (2), we have time and country fixed effects. 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.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 – 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 C7 – 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 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 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 C8 – 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 net national saving as a share of GDP, the level of the real long-term interest rates and the variation of unemployment. The sample includes 20 OECD countries 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 C9 – House price and population age structure

	Benchmark	Reduced form		
	(1)	(2)	(3)	(4)
	All sample	All sample	Low population flow	High population flow
Constant	−1.99***	−1.67***	−1.53	−1.63*
	(0.54)	(0.49)	(1.06)	(0.83)
Variation of population between 1980 and 2014				
0-19	0.35	0.59	1.35*	−0.23
	(0.52)	(0.52)	(0.64)	(0.91)
20 - 49	0.54			
	(0.32)			
50 plus	1.55***			
	(0.25)			
Variation of population between 1960 and 1994				
0 - 29		0.23	−0.02	0.37
		(0.31)	(0.48)	(0.53)
30 plus		1.88***	1.73*	1.62*
		(0.51)	(0.90)	(0.75)
Controls				
ΔReal Disposable income	0.14	0.21	0.45**	0.09
	(0.18)	(0.17)	(0.19)	(0.17)
Real long term interest rate	0.04	0.03	0.06	0.02
	(0.03)	(0.03)	(0.14)	(0.03)
ΔCredit	0.14***	0.15***	0.12**	0.16***
	(0.03)	(0.03)	(0.04)	(0.04)
ΔUnemployment	−2.42***	−2.36***	−2.72***	−2.08***
	(0.37)	(0.35)	(0.56)	(0.40)
Observations	646	646	340	306
No. of countries	19	19	10	9
R^2 within	0.34	0.33	0.41	0.31

Note: Panel linear regression of the growth ofthe real housing price index in percentage, on the growth in percentage of the 0-19 years old, the 20-49 years old, the 50 years old and above or 30 years old and above population between 1960 and 1994, controlled by the growth in percentage of the index of real disposable income, the growth of the credit to the private non financial sector, the level of the real long-term interest rates and the variation of unemployment. The sample includes 19 OECD countries 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 C10 – 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.

Table C11 – Residential investment, migration and natural balances

	(1) All sample	(2) Low population flow	(3) High population flow
Constant	5.63***	4.49***	6.80***
	(0.60)	(0.80)	(0.39)
Migration balance			
20-49	0.69***	0.01	0.94***
	(0.19)	(0.25)	(0.18)
Natural balance			
20-49	1.01***	1.36***	0.90***
	(0.23)	(0.34)	(0.21)
Population growth			
0 - 19	0.35*	0.53***	0.44
	(0.17)	(0.16)	(0.24)
50 plus	0.04	0.69	−0.60*
	(0.41)	(0.46)	(0.32)
Controls			
Real House price growth	0.04***	0.05***	0.03**
	(0.01)	(0.01)	(0.01)
Real Disposable income growth	0.01	0.05	−0.00
	(0.03)	(0.03)	(0.05)
Real long term interest rate	−0.04***	−0.03	−0.03***
	(0.00)	(0.04)	(0.00)
ΔUnemployment	−0.17	0.02	−0.28*
	(0.11)	(0.11)	(0.15)
Observations	560	280	280
No. of countries	20	10	10
R^2 within	0.38	0.34	0.48

Note: Panel linear regression of the percentage share of GFCF dwelling on GDP, both in volumes, on 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.

Documents de Travail

570. G. Verdugo, "Real Wage Cyclicality in the Eurozone before and during the Great Recession: Evidence from micro data" September 2015

571. M. Bussière, L. Ferrara and J. Milovich, "Explaining the Recent Slump in Investment: the Role of Expected Demand and Uncertainty" September 2015

572. M. Albert, C. Jude and C. Rebillard, "The Long Landing Scenario: Rebalancing from Overinvestment and Excessive Credit Growth. Implications for Potential Growth in China" October 2015

573. P. Andrade, G. Gaballo, E. Mengus and B. Mojon, "Forward Guidance and Heterogeneous Beliefs" October 2015

574. C. Jadeau, E. Jousselin, S. Roux and G. Verdugo, "The behaviour of French Firms during the Crisis: Evidence from the Wage Dynamics Network Survey" October 2015

575. C. Berson and N. Ferrari, "Financial incentives and labor market duality" October 2015

576. K. Istrefi and B. Vonnak, "Delayed Overshooting Puzzle in Structural Vector Autoregression Models" October 2015

577. S. Avouyi-Dovi and J-G. Sahuc, "On the sources of macroeconomic stability in the euro area" November 2015

578. F. Canova, F. Ferroni and C. Matthes, "Approximating time varying structural models with time invariant structures" December 2015

579. R. Cezar and O. R. Escobar, "Institutional distance and foreign direct investment" December 2015

580. J. Carluccio, A. Cuñat, H. Fadinger and C. Fons-Rosen, "Offshoring and Skill-upgrading in French Manufacturing: A Heckscher-Ohlin-Melitz View" December 2015

581. G. Cette, R. Lecat et A. Ould Ahmed Jiddou, "Margin rate and the cycle: the role of trade openness" February 2016

582. G. Horny, S. Manganelli and B. Mojon, "Measuring Financial Fragmentation in the Euro Area Corporate Bond Market" February 2016

583. S. Osotimehin and F. Pappadà, "Credit frictions and the cleansing effect of recessions" March 2016

584. F. Langot and A. Pizzo, "Accounting for Labor Gaps" March 2016

585. P. Fève and J.-G. Sahuc, "In Search of the Transmission Mechanism of Fiscal Policy in the Euro Area" March 2016

586. G. Cette, J. Fernald and B. Mojon, "The Pre-Great Recession Slowdown in Productivity" March 2016

587. E. Gautier, D. Fougère and S. Roux, "The Impact of the National Minimum Wage on Industry-Level Wage Bargaining in France" March 2016

588. A. Bergeaud, G. Cette and R. Lecat, "The role of production factor quality and technology diffusion in 20th century productivity growth" April 2016

589. A. Carriero, S. Mouabbi and E. Vangelista, "UK term structure decompositions at the zero lower bound" April 2016

590. M. Girotti, "How Monetary Policy Changes Bank Liability Structure and Funding Cost" April 2016

591. C.Monnet, C.Wolf, "Demographic Cycle, Migration and Housing Investment : a Causal Examination" April 2016

Pour accéder à la liste complète des Documents de Travail publiés par la Banque de France veuillez consulter le site : www.banque-france.fr

For a complete list of Working Papers published by the Banque de France, please visit the website: www.banque-france.fr

Pour tous commentaires ou demandes sur les Documents de Travail, contacter la bibliothèque de la Direction Générale des Études et des Relations Internationales à l'adresse suivante :

For any comment or enquiries on the Working Papers, contact the library of the Directorate General Economics and International Relations at the following address :

BANQUE DE FRANCE
49- 1404 Labolog
75049 Paris Cedex 01
tél : 0033 (0)1 42 97 77 24 ou 01 42 92 63 40 ou 48 90 ou 69 81
email : 1404-ut@banque-france.fr