First-time house buying and catch up: A cohort study

RENATA BOTTAZZI[a], THOMAS F. CROSSLEY[b] AND MATTHEW WAKEFIELD[c]

a. *University of Bologna and Institute for Fiscal Studies.* renata.bottazzi@unibo.it.
b. *University of Essex and Institute for Fiscal Studies.* tcross@essex.ac.uk
c. *University of Bologna and Institute for Fiscal Studies.* matthew.wakefield@unibo.it

April 2015*

Abstract

We investigate whether different ages of first-time house buying lead to persistent differences in homeownership between cohorts. Our data span nearly forty years and multiple cycles of England's volatile house prices. Ownership rates at thirty have differed substantially, with a significant negative association with prices. The persistence of differences is assessed using synthetic cohort techniques. Two methods of dealing with measurement error problems both indicate that cohorts with low ownership rates at thirty catch up almost all of the ownership gap by forty. Earlier access to homeownership may result in the ownership of slightly larger homes at around forty.

JEL Classification numbers: R21, R31

Keywords: homeownership, synthetic cohort data, measurement error.

I. INTRODUCTION

The birth cohort born in 1967 turned twenty-two in 1989. Some were graduating from university, while others had been in the labour market a few years. Most aspired to starting families and owning their own homes. In the United Kingdom, these twenty-two year olds faced a housing market in which average prices had been rising for seven years, and had risen 70% in real terms in the last four years. The ratio of average house prices to average earnings was 5.5. By contrast, when the cohort of 1975 turned twenty-two in 1997, house prices were more than 20% lower than in 1989. Incomes had been catching up with prices, so that the house price to earnings ratio was 4. In short, this cohort faced a very different housing market in early adulthood than the cohort that turned twenty-two eight years earlier.[1] As widely noted in the press, the rapid recovery of the housing market over the last few years has generated another cohort facing a booming housing market in early adulthood.[2] Do these differences matter?

These differences may matter both in the short run and in the long run. In the short run, cohorts faced with difficult housing market conditions may, on average, be delayed in 'getting on the property ladder.' Perhaps even more seriously, these differences may also matter for the longer run homeownership rates of a cohort, and for other outcomes of those cohorts. Some members of a cohort that is delayed in its initial ownership transitions may find that they are never able to make the transition to owning their own home, and the ownership rate of the cohort may never 'catch up' to that of cohorts that faced more favourable initial conditions. Additionally or alternatively, delayed ownership may be associated with consuming less housing (owning smaller houses) at later ages.

Literature suggests that factors that delay or discourage ownership may be associated with long-run changes in other outcomes, and so in well-being. Recent research in the U.S. suggests that house prices can affect fertility decisions (Dettling and Kearney, 2014) and that home ownership and home equity are the major drivers of cohort-level differences in wealth

at retirement (Lusardi and Mitchell, 2007). Research has also shown that home owners use home-equity withdrawal as a financial buffer and smoothing mechanism in both the U.S. and U.K. (Hurst and Stafford, 2004; Benito, 2009); this consumption smoothing mechanism is not available to non-owners. Finally Banks et al, (2010) demonstrate that earlier ownership is an important hedge against house-price volatility in the U.K.

Given the evidence of links between home-ownership and outcomes that affect well-being, it is important to measure how factors such as housing market conditions are associated with differences in home-ownership. Recent theoretical modeling (Bottazzi, Low and Wakefield, 2007) suggests that housing market conditions early in a cohort's housing career matter in the short run, but not in the long run. Simulations indicate that disadvantaged cohorts catch up, so that they have comparable homeownership rates as they approach their fifties. This would significantly diminish the role of early life housing-market conditions in shaping later outcomes. However, these simulation results are sensitive to modeling choices, and so an empirical assessment of these questions remains important.

In this paper we provide such an empirical assessment. We use the repeated cross sections of the Family Expenditure Survey/Expenditure and Food Survey (FES/EFS) from 1969 to 2007 to answer several questions.[3] First, as each birth cohort reaches adulthood, how do their transitions to homeownership vary with general housing market conditions? Specifically, if we compare two cohorts, one that faced a property boom in their twenties and one facing a property slump, how different are their ownership rates at age thirty? Second, how persistent are the resulting differences? That is, do the homeownership rates of these two cohorts converge at older ages? Finally, are age thirty differences in ownership associated with other aspects of housing tenure at older ages, including house size, and mortgage status?

The FES/EFS is not a true panel, in that individuals (or individual households) are not followed over time. However, because the FES/EFS surveys provide us with a representative

sample of the population in any survey year, it also gives us a representative sample of each birth cohort in any survey year. We use the repeated cross-sections of the FES/EFS in two ways. First, we hold age constant and look at variation in ownership across birth cohorts. The resulting age-specific ownership rate evolves over time as different cohorts pass through the age of interest, but will differ from the overall homeownership time-series, which is at every point in time an amalgam of different cohorts and ages. Second, we employ synthetic-cohort methods, holding cohort constant and using the repeated surveys to track birth cohorts across ages (and hence over time.) This allows us to study whether cohorts catch up in homeownership, which is a novel contribution of the paper.

A brief preview of our results is as follows. First, over the past forty years there has been considerable cross-cohort variation in the rate at which different birth cohorts transition to homeownership. Ownership rates at thirty have ranged from around fifty percent to approximately seventy percent. Second, this variation is correlated with house price developments over time, although that relationship seems stronger before 1990 than since. Overall, our results suggest that when a birth cohort faces a house price to average income ratio at age thirty that is one standard deviation above trend, then the homeownership rate of that birth cohort at age thirty is approximately 4 percentage points lower. Third, there is strong negative correlation between cohort ownership rates at age thirty and subsequent growth in ownership: cohorts that have low ownership at thirty appear to have fast growth in homeownership subsequently. Historically, cohorts with low homeownership rates at thirty have closed about 80% of the "ownership-gap" by the time they reach age forty. Finally, we find evidence that cohorts that get on the housing ladder earlier own larger homes in middle age, but no evidence that they are more likely to own their homes outright, or have lower mortgage payment to income ratios.

The rest of the paper is organized as follows. Section II elaborates on the historical context for our study, and describes the data and methods we employ (with further details of

our data and methods provided in an online only Technical Appendix). Section III then analyses the pattern of ownership at age thirty across birth cohorts and how this correlates with house price developments. Section IV considers the question of the persistence of these differences and Section V considers the association of age-30 ownership rates with later life housing outcomes. Section VI concludes.

II. CONTEXT, DATA, AND METHODS

Context

This study concerns the rate at which households have been able to get onto the housing ladder during the last forty years. Housing market conditions, most notably house prices, have affected the affordability of homeownership during this period. Additionally, trends in access to finance and public policy reforms have affected the accessibility of home purchase. In this subsection we describe trends in average house prices and in their relationship to average family incomes,[4] and in credit conditions, and also outline an important public policy programme that has affected homeownership rates (at different ages) in the years of the study. These trends and changes are factors that we will exploit, or need to take account of, in the analysis of the later sections of the paper.

Over the 35 years to 2007, England experienced three house price booms and two periods of significant house price decline. This can be seen in the "House Price" panel of Figure 1, which shows a quarterly measure of the (mix-adjusted) average real house price (dashed black line) and the (mix-adjusted) average real house price for first-time buyers (solid black line), for England, superimposed over a measure of the cyclical element of national GDP. Trends and cyclical movements are seen to be remarkably similar for all properties and for properties bought by first-time buyers, with first-time buyers buying properties that cost around 20 to 25 per cent less than the overall average. Over the whole period 1969-2007, average real house prices (and first-time buyer house prices) in England

increased by a multiple of almost four. As mentioned, this did not happen through a continuous upward trend. House price booms are seen in the early and middle 1970s, in the second half of the 1980s (during which period average real house prices rose by over 60% in four years), and in the period between 1995 and the early 2000s (until 2007). Real terms house price falls were experienced between 1974 and 1977 (a period which was not followed by sustained price growth until after 1985), and in the first half of the 1990s (during which period average real prices fell by almost forty percent). By superimposing house prices over deviations of GDP from trend we can see that during the 1970s and 1980s house price cycles tended to coincide with movements in the underlying economy. However, after the recession of the early 1990s and a trough in house prices in the mid 1990s, GDP has not deviated strongly from trend while house prices displayed strong growth for more than 10 years.

[Figure 1 about here]

The right hand "House Price to Income" panel of Figure 1 displays, as a solid black line, the ratio of (average) first-time buyer house prices to average incomes of those aged around 30 (again superimposed over deviations of GDP from trend). This ratio might be thought a stronger indicator of housing "affordability" than the price level, and is an important variable in our regression analyses later in the paper. With the possible exception of the final run up in the price level after the mid-1990s, using the price to income variable largely takes out the upward drift over time seen in house prices, as average incomes have also grown. However, house price cycles have been more pronounced than cycles in average family incomes and so the price to income ratio shows cyclical movements that reflect the price cycles discussed in the previous paragraph: the price to income ratio spikes up in the early 1970s, in the late 1980s, and again after the middle 1990s, and the ratio fell in the middle 1970s and in the early to middle 1990s.

Changes in house prices are not the only factor that have changed and will have affected the ability of households to get on and climb the property ladder during the last four

decades. This is also a period during which substantial changes in credit markets took place. In addition, some public policy changes have been important.

Regarding credit conditions, the 1980s was a period of substantial credit market liberalization. As is clear from the important work of Fernandez-Corugedo and Muellbauer (2006), detailed modelling work is needed to accurately quantify credit conditions.[5] Their exercises provide a clear indication of the extent to which credit conditions changed during the 1980s. Describing their two measures of consumer credit conditions between 1975 and 2001, they write that "[b]oth indices increase in the 1980s, peaking towards the end of the decade. They fall partway back in the early 1990s, before increasing again towards the end of the sample" (ibid, p.4). A close look at their indices shows that the increase during the 1980s was particularly rapid during the first three years of that decade.

At the same time as the financial liberalization was taking hold, a major policy reform was also affecting the English housing market. This was the "right to buy" scheme which allowed council tenants (i.e. those renting social housing) the right to buy their properties at prices that were discounted compared to market values, with discounts depending on the length of tenancy. This became national policy[6] with the passing of the Housing Act of (October) 1980, and resulted in a transfer of households from the social renting sector into owner-occupation. Figure 2 shows official statistics for the number of right to buy sales of local authority properties in England for each (financial) year from 1980/81.[7] We see that there were particularly big spikes in such house sales at either end of the 1980s, with a smaller peak in the early 2000s; by 2008/09 almost 1.8 million local authority properties had been sold.[8]

[Figure 2 about here]

Data

Our aim is to study the multiple housing booms and busts that occurred over the last four decades in England. Panel data that track the same individuals over this entire period do not exist. The British Household Panel Study, for example, has excellent data on housing arrangements, but begins in 1991. Thus, only fifteen birth cohorts can be observed at any age, and only one house-price boom can be studied. While much important housing research can be done with these data, such as about the decision to leave the parental home (Ermisch, 1999), it is of limited use for our purposes. Instead, we use the Family Expenditure Survey/Expenditure and Food Survey (FES/EFS) which is available since 1968 and therefore allows multiple comparisons between cohorts that experienced favourable and unfavourable housing market conditions in their late twenties.

The FES/EFS is an annual cross section of around 7,000 households, who record a two-week diary of their spending and information about purchases of durables and/or expensive items in recent months prior to the interview. Importantly for our study, the survey provides information on the housing tenure of respondents, on the number of rooms in their house, on their mortgage debt, as well as on their income, education, and family structure. In all our calculations we use the appropriate survey weights.

We supplement the FES/EFS data with data on house prices and on sales of local authority housing through the right to buy scheme, a measure of interest rates and a credit conditions index. We use official Government national and (for house prices) regional data, provided through the Office for National Statistics (ONS)[9]; detailed information on how the house-price data are set up may be found in the online Technical Appendix. The data on right to buy sales are those underlying Figure 2 above. The interest rate that we use is the 90 day Treasury Bill Discount Rate[10]. The Credit Conditions Index that we use is the unified measure generated by Fernandez-Corugedo and Muellbauer; it is discussed at more length in our online technical appendix and in detail in Fernandez-Corugedo and Muellbauer (2006).

Methods

We use the repeated cross-sections of the FES/EFS in two ways. First, we hold age constant and look at variation in ownership across birth cohorts. The resulting age-specific ownership rate evolves over time as different cohorts pass through the age of interest, but will differ from the overall homeownership time-series, which is at every point in time a size-weighted average amalgam of different age groups (and hence birth cohorts). The repeated cross-sections allow us to focus on each cohort as it passes through young adulthood.

Second, the FES/EFS allows us to study the housing careers of more than thirty birth cohorts through synthetic cohort analysis. The basic idea of synthetic cohort analysis is as follows. With repeated cross sections we cannot track individuals over time. However, in each survey year we get a representative sample from each birth cohort, and so by using successive cross sections, we can follow the average characteristics of a birth cohort through time. In particular, for any birth cohort, we can estimate its ownership rate in every survey year and hence at different ages. Myers (1999, 2001) has emphasized the importance of accounting for cohort effects in the analysis of housing careers, and the utility of cohort studies as an important alternative to cross-sectional and longitudinal approaches to exploring housing patterns. At the same time, cohort analysis often offers a longer time span than does the available panel data (as is the case here); in addition, the use of repeated cross-sections to follow synthetic cohorts avoids the attrition and small sample problems that often limit panel data analyses.

With thirty-nine FES/EFS surveys available to us (1969-2007) we can potentially follow some cohorts for thirty-nine years. However, we largely focus on ages thirty to forty (and sometimes ages thirty to fifty).

Although the FES/EFS are household surveys, we believe the appropriate unit of analysis is the individual and in this study we follow cohorts of individuals. Although it takes some care, birth cohorts of individuals can be constructed from the FES/EFS. The concept of

a “household life-cycle” is commonplace in economic studies, but such an approach has several drawbacks. A household is a collection of individuals, each of whom may belong to different cohorts and, at any given time, may be at a different stage of the life-cycle. Although the ambiguity of a “household life-cycle” is well recognized, this ambiguity is often ignored because of the potential complexity of discerning individual profiles of household members from household data.

Moreover, many transitions in housing arrangements are associated with household formation or dissolution or with changes in household composition. Recent NHPAU research on affordability has focussed on the issue of household formation (NHPAU, 2008). Myers (1990) explains this concern with following the housing careers of households: “[w]hereas most housing research begins with the behaviour of households, the logical prior concern in this type of research is with the formation of households from a population” (p. 14). Housing studies that followed housing choices of (cohorts of) couples would miss much of the important action.

Instead, in this study, we follow cohorts of individuals. To generate individual birth cohorts from household data, we create individual observations whenever we see an individual of a certain age and gender in a household record. The FES/EFS contains information on household and individual characteristics thus allowing us to create detailed records from which to construct individual birth cohorts for adults of all ages. Hence we are able to track changes in housing tenure alongside changes in family composition for both men and women.

When structuring the data into cohorts of individuals, some care is required with allocating homeownership. We take ownership to be a shared state, so that if we see a couple living in a property that is owned (with a mortgage or outright) by either member of the couple, then our data records both members of the couple as being owner-occupiers. Thus when we consider counts of individuals, both of these individuals will be counted as owners.

However, we do not necessarily allocate the same ownership status to all members of a household. In particular, we are careful about how we allocate ownership for young adults who are still living in the parental home. Such individuals will appear as observations in our dataset, which includes all adults. However, even if the data record that the parents own their home, our analysis does not treat the children as home-owners. Recording ownership state in this way ensures that there is not an apparent fall in ownership in the early and middle twenties as individuals move out of home (often into the rental sector), followed by an increase when the same individuals become (first-time) buyers.

In presenting our analyses, we sometimes write as if the increase in the homeownership rate for a given cohort is the proportion of that group that became home owners between one year and the next. That is, we interpret this change as the gross flow in to housing between one year and the next. However, the flow that we observe is actually the net flow. That is, it is the number moving into homeownership, net of the number transiting in the other direction back in to the rental sector. Whether this net flow provides a good approximation to the gross flow depends on how the number of individuals buying houses compares to the number of individuals in the same group (of the same age) who move from being owners back into the rental sector.

We undertook some preliminary analysis of this issue using the BHPS. Because the BHPS is a true panel, both gross and net flows are observed directly. Fortunately, for individuals in the age ranges that we are considering, net flows approximate gross flows quite closely. Among individuals in their twenties, there are relatively few individuals transiting back into renting because relatively few already own. Around age thirty, when the ownership rate is around 60% (cf. Table 2), the proportion of owners that switch to renting is around 2%, and this tends to decline with age throughout the working life (being around 1% at age forty). Thus, though home-owners are in the majority at these ages, the numbers switching back to rental remain small.[11] Thus, while the reader should bear in mind that what we

actually observe are net flows in to ownership, it is not too large a distortion to discuss our results as if we were observing a gross flow.

The size of the FES/EFS dataset allows us to split our analyses by region. Regional analysis is of independent interest; moreover, splitting by region potentially provides additional variation in prices to exploit. However, the synthetic cohort part of our analysis rests on the assumption that the composition of the cohort being followed is fixed over time. This assumption might be undermined if migration flows between regions are sufficiently large. We have investigated this issue empirically and concluded that it is reasonable to follow cohorts defined by birth-year and region. Further details are provided in the online Technical Appendix.

III. GETTING ON THE HOUSING LADDER: HOMEOWNERSHIP AT THIRTY

We begin by calculating the homeownership rate of individuals aged twenty-nine to thirty-one for every year in the data. Figure 3 displays the homeownership rate for thirty-year olds through time. For the figure, the ownership rate was calculated using survey weights so that resulting rate should be representative of the population of interest. The ownership rate is displayed alongside the ratio of the first-time buyer house price to average incomes around age 30 (the same price to income ratio that was displayed in the right-hand panel of Figure 1, but at an annual frequency). The house price series is in grey and dashed with data points marked by crosses, and the house price boom that began in the mid-1990s is clear in the graphs. The ownership series is a solid black line with rings as markers. Note that because we are holding age constant, the x-axis measures both survey year and birth cohort: thirty year-olds in 1970 are from the 1940 birth cohort and so on.

Over the past forty years there has been considerable cross-cohort variation in the rate at which different birth cohorts have transited to homeownership. Figure 3 shows that

ownership rates at thirty range from around 50 percent to approximately 70 percent, with both up and down swings.

[Figure 3 about here]

The data in Figure 3 do suggest a relationship such that high prices (relative to income) coincide with low ownership among thirty-year olds, particularly before 1990. The peaks and troughs in prices before the mid-1980s approximately correspond to troughs and peaks in the age-thirty ownership rate. It is also the case that the strong run up in the house price after 1995 is associated with a downward drift in age-thirty ownership (although this downward drift did begin before house prices began to climb). However, between 1980 and 1985 the noticeable feature of the data is a strong surge in the age-thirty ownership rate, from around fifty-five percent, past its previous peak of almost sixty percent, and up to almost seventy percent. While this increase seemed to reverse somewhat as house prices began to grow rapidly in 1986 and 1987, it is worth noting that the reverse began before prices reached their peak in the late 1980s, but, as already noted, turned to a secular decline even while prices were falling at the beginning of the 1990s. It is likely that pressures other than prices – such as the already noted credit liberalization, and the "right to buy" policy – were affecting ownership rates strongly at some points between 1980 and the early 1990s, and in the figure this swamps the effect of the prices relative to income on affordability.

Disaggregations illustrate that trends in the home-ownership rates of thirty year-old men and women, or of thirty year-olds in broad regions (North, Midlands and South) of England, have been similar to the patterns shown in Figure 3.[12] Figure 4 explores differences or similarities in trends for single individuals and individuals that live as part of a couple. Specifically, the figure again shows the ownership rate of all thirty year olds, and also shows the ownership rates for the subsets of individuals that are single and that live as part of a couple, and finally the proportion of thirty year-olds that are in a couple. The figure thus shows differences in ownership rates between singles and couples, and also highlights that

trends in the overall ownership rate have been related not only to trends in ownership among singles and couples, but also to changes in the proportion of the population who are living as part of a couple.

[Figure 4 about here]

A simple accounting identity is that:

Ownership rate of 30 year-olds =

Fraction in a couple x ownership rate among coupled 30 year-olds

+ Fraction single x ownership rate among single 30 year olds

Figure 4 shows that the ownership rate among coupled thirty-year olds has fallen relatively little since 1990, while the ownership rate of single thirty-year olds has fallen a bit more. Young coupled individuals have, unsurprisingly, always had higher ownership rates than young singles, and the proportion of thirty-year olds in couples has been falling (from over eighty percent in the 1970s, to around two-thirds in the early 1990s). This accounts for a substantial component of the fall in the overall ownership rate of thirty-year olds.

While this observation provides a mechanical explanation of how the recent decline in ownership at thirty has occurred, a causal inference should not be drawn. It could be, for example, that the decline in the fraction of thirty year-olds who are a member of couple has been driven by a declining affordability of homeownership. The direction of causation is unclear.

To quantify the relationship between house prices and ownership at thirty observed in these figures, we now turn to probit models for homeownership. Probit or logit models for homeownership (tenure choice) estimated on micro-data are well known in the literature. For example, Linneman and Wachter (1989) estimate Logit models for homeownership on American microdata, with a particular focus on wealth constraints. Hilber and Liu (2008) also

estimate a logit mode of tenure on American microdata but focusing on the roles of own and parental wealth, and location preference in explaining the black-white ownership gap in America. Bourassa and Hoesli, (2010) estimate a logistic regression for tenure choice on Swiss microdata focusing on wealth constraints and the relative costs of owning and renting.

Our analysis is differentiated from these papers and the related literature by our use of data spanning a much longer period (and therefore, affording much more temporal variation in prices), and by our focus particularly on ownership at age thirty. The papers cited above pool households of a range of ages. They do typically include age variables among the demographic controls in their tenure choice models. But the linear index models they use constrain the estimated relationship between tenure choice and other variables (for example prices) to evolve with age in a very particular way. By focusing very narrowly on young adults, we can estimate the correlates of ownership at age thirty without constraining the parameters to fit the choices of other age groups. Against this, a number of the papers cited above model the relative costs of ownership and rent, or the wealth and credit constraints, in more detail than we do here. Like Hilber and Liu, we take a reduced form approach.

Estimates of four probit models of homeownership at age 30 are reported in Table 1. All are based on pooled data on individuals from many waves of the FES/EFS. In all three cases the dependent variable is whether or not the individual owns a property at age thirty. They differ by the set of conditioning variables and the range of years of data employed. The latter is driven by the former: in each case we use all the years for which data on the included conditioning variables are available.[13] As always we make appropriate use of survey weights. Standard errors are clustered on the year level to reflect the fact that the most aggregated regressors in our specifications vary at the annual level.

The first three models contain the ratio of house prices to income as a measure of housing affordability. For this purpose, house prices are taken from the (mix-adjusted) first-time buyers series at region and quarter level. Income is averaged among 26-34 year-olds (in

region and quarter). The regressions also include the interest rate as a measure of the cost of homeownership.

The models also include two variables measuring the number of "right to buy" sales of local authority housing in England. These two variables are intended to capture the impact of right to buy on the number of properties up for sale in a particular year, and on the size of the stock of properties in the owner-occupied sector, and respectively measure the number of right to buy sales (in hundreds of thousands) in the (financial) year in which an individual is observed and the cumulative number of sales (again in hundreds of thousands) since the right to buy became a national scheme in late 1980. Given the years in which the right to buy has been an active policy, it is possible that our right to buy variables also pick up some effects of the financial market changes discussed in the first part of Section II.[14]

Additional controls that are common to all the models include the gender of the individual and whether he or she is a member of a couple; the log of family income; and the number of children.

The first model (in the left-most column) is our base specification and it is estimated on pooled data from 1969 to 2007. The second model (middle-left column) adds a measure of the individual's education and is estimated on data from 1978 on (when education began to be recorded in the FES/EFS). The third model (middle-right column) adds, in addition to education, the credit conditions index developed by Fernandez-Corugedo and Muellbauer (2006). As we have this variable only until 2005, this model is estimated on pooled data from 1978 to 2005.[15] The fourth model excludes the house price to income variable, but otherwise matches the third model in terms of regressors and data.

The variable for right to buy sales this year is significant in the first and second specifications, suggesting that ownership at thirty has tended to be higher in years with greater numbers of right to buy sales, and the variable recording the cumulative number of

sales is never significant. Unsurprisingly, family income and family characteristics (couple and number of children) are also strongly and significantly associated with homeownership.

The ratio between house price and income is significant and has a negative sign in all three specifications in which it is included, indicating that higher prices, relative to income, are associated with lower ownership rates among thirty year-olds. That is, thirty year-olds own less when housing is less affordable. In the first column, the marginal effect on the price-to-income variable suggests that if the house price to income ratio goes up by one standard deviation – which is almost exactly 1[16] –, then homeownership would be slightly more than 4 percentage points lower. The strength of this relationship is robust to adding controls for education and credit conditions in columns 2 and 3: the relevant marginal effects in all three columns are of very similar magnitude.

Credit conditions (measured by the index we use) are known to have important effects on the level of real house prices in Britain (Cameron, Muellbauer and Murphy, 2006).[17] Thus, one possible explanation for why the credit conditions index is not significant and does not substantially affect the other parameters in the model, is that the effects of credit conditions "act through" house prices. The house price variable was excluded from the model in column 4, in order to investigate this hypothesis. However, while excluding the house price to income ratio does lead to a change in sign for the coefficient on the credit conditions variable (and that on the interest rate), these variables remain insignificant.[18]

We experimented with adding lags of the house price to income variable to the models presented. This allowed us to investigate whether price effects are stronger if prices have been persistently high relative to incomes as a cohort approaches thirty, than if they become high only near age thirty. We did not find significant evidence of such "dynamic price effects". A single lag of anything between one and five years was not significant and did not much affect the coefficient on the current price to income, while a formal statistical test indicated that even adding all five lags together did not significantly improve the

explanatory power of the model. We conclude that the current price to (average) income is a sufficient control for price effects.[19]

The results (notably for the price effects) are robust. Similar results are obtained whether we use the survey weights or not, and when we use data for Great Britain rather than for England alone. House price (to income) effects appear slightly smaller in magnitude (less negative) if the house price variable is based on the variation in the prices of all properties rather than in the prices of properties bought by first-time buyers; this is consistent with the notion that first-time buyer prices are more directly relevant to the house purchase decisions of our sample of interest. Using only national (as opposed to regional) house price variation also leads to a reduction in the magnitude of the point estimate of the marginal effect of the house price to income. Omitting individual characteristics (couple and number of children) from the model makes very little difference to the marginal house price (to income) effect. Adding a post-1980 dummy to the model reported in column 1 does slightly reduce the point estimate of the marginal effect for the house price to income variable (to three percentage points), and if we split the sample by education groups, the lower educated exhibit a stronger association with house prices. However, even in these cases (as in all the cases discussed in this paragraph) we cannot reject the hypothesis that the effect estimated in our baseline model (usually column 1 of Table 1) is equal to the point estimate obtained in the modified model. [20]

[Table 1 about here]

To summarize: homeownership rates at age thirty vary substantially across birth cohorts and the data support the idea that unfavourable housing market conditions in early adulthood are associated with delays in the transition of birth cohorts into homeownership.

IV. IS THERE CATCH UP AFTER THIRTY?

We now turn to the question of whether those cohorts that were less able to get onto the ladder by thirty were nonetheless able to "catch up" with other cohorts at older ages. Do early

differences in the rate of transition to homeownership persist into later life? Are other outcomes, such as the amount of housing assets owned, affected by ownership at age thirty? This is a critical issue from a number of policy perspectives. For example, as noted in the introduction, homeownership is a strong predictor of economic security in retirement.

Figure 5, in which we present homeownership rates at different ages across years, provides a first look at this question. The solid black line is the ownership rate for thirty year-olds, the dashed grey line (square markers) is this rate for forty year-olds, and so on with the solid dark grey line (round markers) being the ownership rate among individuals aged seventy.

We see that for each group there is a substantial increase in the proportion of owners during the period before 1990, a time trend that reflects the right to buy policy and credit market liberalization, among other things. After 1990, the homeownership rate for thirty year olds declines sharply, as we saw for the ownership rate of thirty year olds in Figure 3. This is a contrast to the ownership rates for other age groups, which stayed roughly constant or even continued to increase slowly. This contrast already suggests some catch up: individuals who were thirty in 1990 did not own substantially more when they were forty than was the case for those who were thirty five years later. Thus the higher ownership of the former group at thirty was offset by later transitions into owning for their successor cohort.

[Figure 5 about here]

Table 2 summarizes the variation in ownership rates across cohorts, at different ages. Panel (a) summarizes ownership rates at thirty and forty for those birth cohorts that we observe at *both* thirty and forty. Panel (b) does the same for ownership rates at thirty and fifty for the smaller set of cohorts that we observe at *both* those ages (see also Table A.1 in the online Technical Appendix). The Table indicates that there is less dispersion in ownership rates across birth cohorts at older ages than at younger ages. This is again indicative of "catch up".

[Table 2 about here]

To look more directly at catch up, in Figures 6 and 7, we plot the increase in ownership between age thirty and age forty (in Figure 6) and between age thirty and age fifty (in Figure 7) against ownership at age thirty. In these figures, each point represents a birth-cohort (and the points are labeled by the birth year of the cohort). In each figure there are four panels. The top-left panel displays the relationship for cohorts defined by birth-year only (that is, for all of England). The remaining three panels repeat the analysis separately for each of three regions: North (top-right), Midlands (bottom-left) and South (bottom-right).

[Figures 6 and 7 about here]

Catch up implies a negative relationship: lower ownership at thirty must be associated with a greater subsequent increase and higher ownership at thirty with a lower subsequent increase. This is exactly what we see in Figure 6 (catch up by forty) and Figure 7 (catch up by fifty). The same pattern is observed at the national level and in each region. For example, in the top left panel of Figure 6, we see that the 1940 birth cohort (thirty in 1970) has a low homeownership rate at age thirty of 47% (see also Table A.1 in the online Appendix) but experiences a substantial increase in homeownership – of 16 percentage points – between ages thirty and forty. In contrast, the 1954 birth cohort (thirty in 1984) has a much higher homeownership rate at age thirty of 71% but experiences very little increase in homeownership – just 2 percentage points – between ages thirty and forty.

However, there is a potential problem with these figures. We know that ownership at thirty is measured with error. For each cohort, it is an estimate, based on the representative sample of that birth cohort found in the appropriate year of the FES/EFS. These estimates are naturally subject to sampling error, and this sampling error is effectively a kind of measurement error (Deaton, 1985). The ownership rate at thirty will be slightly over-estimated for some cohorts, and slightly under estimated for others. This measurement error may affect the figures in two ways.

First, ownership at thirty appears on the horizontal axis in each figure. Measurement error in the horizontal variable creates attenuation bias and makes the relationship appear flatter than it actually is; in the case of a negative relationship like that documented in Figures 6 and Figure 7, this means the slope is *less* negative than it should be, understating the true degree of catch up.

Second, ownership at thirty is also used to construct the variable (change in ownership) on the vertical axis, which it enters negatively. Ownership rates at forty will be subject to sampling error as well, but because these are based (for each cohort) on an independent sample, the sample errors in ownership at forty will be unrelated to sampling errors in ownership at thirty. The way ownership at thirty features on both axes means that cohorts that have positive measurement errors in ownership at thirty will appear to have smaller subsequent increases in ownership. Measurement error in ownership at thirty therefore creates a spurious negative correlation between change in ownership (on the vertical axis) and ownership at thirty (on the horizontal axis.) This makes the relationship appear more negative than it actually is, overstating the true degree of catch up. (These arguments are formalized below).

These two effects operate in opposite directions so that the direction of net bias is unclear. One might hope that they roughly cancel, but there is no guarantee that this is the case. Therefore, we next employ two methods that allow us to circumvent these measurement problems and quantify the degree of catch up.[21]

To quantify the catch up suggested by the figures we regress the change in the ownership rate between thirty and forty, on the ownership rate at age thirty. Catch up implies a negative coefficient on the initial condition (ownership at age thirty). If subsequent increases in homeownership are unrelated to ownership rates at thirty, then the coefficient on the latter should be zero. Complete catch up corresponds to a coefficient of minus one. In this

case of complete catch up, cohort ownership rates at forty are not predicted by ownership at thirty.

We focus on catch up between thirty and forty because this maximizes the number of birth-cohort observations we can use in estimation. (In a given set of survey years, not all birth-cohorts are seen at all ages, and more cohorts are observed at both ages thirty and forty than at ages thirty and fifty.) Note that, unlike the econometric model reported in Section III, which was estimated on individual level data, the model here is estimated on cohort -level data (each observation is a birth cohort of individuals). It is infeasible to estimate this growth model on the pooled individual data because each individual is observed only once: it is only the birth cohort that is observed at more than one age. We use a linear model this time because our dependent variable is not dichotomous but rather is measured in percentage points. The results are presented in Table 3.

[Table 3 about here]

We first estimate this model by ordinary least squares and the results of this estimation procedure are presented in column (1). In this regression the ownership rate at thirty is significant and negative – as we would expect given the figures in the previous subsection. The coefficient of -0.871 suggests that around 87% of the variation in birth cohort homeownership rates at age thirty is made up by age forty, and we cannot reject a coefficient of -1 (i.e. complete catch up). In column (2) we add to this model a dummy for reaching thirty in or after 1981, and the fraction of the cohort that were in a couple at thirty.[22] This results in a slightly larger estimate of the extent of catch up of about 95%.

However, these ordinary least squares estimates suffer from exactly the same problem as was described for Figures 6 and 7. Biases arising from measurement error in ownership at age thirty may lead to either over- or underestimates of the degree of catch up.

There are two possible approaches to overcoming these problems. The first approach is to re-estimate our regression model by two-stage least squares (2SLS), using the ownership

rate at age twenty-nine as an instrument for our mis-measured independent variable, ownership at thirty. The ownership rate of a birth-cohort at age twenty-nine is very closely related to its ownership rate at thirty. The ownership rate at twenty-nine is measured with error, for the same reasons that the ownership rate at thirty is. However, because for each cohort the ownership rates at twenty-nine and thirty are based on different survey years (and hence independent samples), the measurement (or sampling) error in the ownership rate at twenty-nine should be unrelated to the measurement (or sampling) error in the ownership rate at thirty (and forty). Thus the ownership rate at age twenty-nine is an ideal instrumental variable in this context.

The results of this exercise are presented in columns (3) and (4) of Table 3. Relative to the OLS estimates, the point estimates of the catch up coefficient are somewhat diminished in magnitude. For example, for the specification with no additional controls, the coefficient goes from -0.871 to -0.835 (so that estimated degree of catch up goes from 87% to 84%). This suggests the net bias in the OLS estimate from measurement error is a small exaggeration (in magnitude). The small size of the net bias is consistent either with the two oppositely-signed biases described above both being small, or of very similar magnitudes (and so offsetting). The key point is that the 2SLS estimate still suggests substantial catch up, although the coefficient is now much less precisely estimated.

The second approach to overcome the measurement error bias is based on the observation that the coefficient of interest can be corrected for the measurement error if an estimate of the degree of measurement error (σ_u^2 below) is available. In the case of pseudo-panel analysis, as observed by Deaton (1985), such an estimate is available because the measurement error is just sampling error of the cohort-year cell mean, which can be estimated by standard methods.

Deaton's corrected estimators do not apply directly to our catch up regressions, but the formulas from that paper can easily be extended to our specification. In our case, the

explanatory variable of interest (x^*) is ownership at thirty, while the outcome of interest ($y^* - x^*$) is the change in ownership between thirty and forty. For the univariate case, let x^* and y^* indicate true variables and let x and y be the variables we observe with errors u and v, respectively (omitting year subscripts for y, x, and ε):

$$y^* - x^* = \beta x^* + \varepsilon$$
$$y = y^* + v$$
$$x = x^* + u$$

We make standard assumptions about the structure of the model and of measurement errors:

$$\text{plim}\frac{1}{n}\sum x^*\varepsilon = \text{plim}\frac{1}{n}\sum x^*u = \text{plim}\frac{1}{n}\sum x^*v = 0$$
$$\text{plim}\frac{1}{n}\sum u\varepsilon = \text{plim}\frac{1}{n}\sum v\varepsilon = 0$$

We denote variances and covariances of true variables and measurement error as

$$\text{plim}\frac{1}{n}\sum x^{*2} = \sigma_{x^*}^2 \qquad \text{plim}\frac{1}{n}\sum y^*x^* = \sigma_{y^*x^*}$$
$$\text{plim}\frac{1}{n}\sum u^2 = \sigma_u^2 \qquad \text{plim}\frac{1}{n}\sum uv = \sigma_{uv}$$

In our case u and v are sampling errors from different independent samples so from now on we assume $\sigma_{uv} = 0$.

The OLS estimator based on observed variables is (with sums taken over years)

$$\beta^{OLS} = \left(\sum x^2\right)^{-1}\sum (y - x)x$$

(Note that throughout this Section, and in our implementation of the estimator, x and y variables are in deviations from means.)

This estimator is not consistent, having

$$\text{plim}\ \beta^{OLS} = \frac{\beta\sigma_{x^*}^2 - \sigma_u^2}{\sigma_{x^*}^2 + \sigma_u^2} = \frac{\beta\sigma_{x^*}^2}{\sigma_{x^*}^2 + \sigma_u^2} - \frac{\sigma_u^2}{\sigma_{x^*}^2 + \sigma_u^2}$$

This last expression characterises the effects of the measurement error. The first element in the last sum above is standard attenuation bias due to measurement error in the x variable, and tends to make the estimator smaller in magnitude. The second element will tend to bias the estimator towards -1 (as the variance in the true x goes to zero).

We can (and do, see column (5) of Table 3) correct the estimate using the following expression:

$$\beta^{FLS} = \left(\sum (x^2 - s_u^2)\right)^{-1} \sum \left((y - x)x + s_u^2\right)$$

where s_u^2 is a consistent estimate of σ_u^2. In our case, this is the variance of the cohort-year sample mean of ownership at thirty, which can be estimated from the pooled micro-data underlying the pseudo panel. In doing this, we weight cohort-year cells to allow for differences in cell size.

Maintaining notation as much as possible (vectors in bold and *t* an index for year), the multivariate case is

$$y_t^* - x_{1t}^* = \boldsymbol{x_t^{*\prime}\beta} + \varepsilon_t$$

with x_{1t} being the first element of the $\boldsymbol{x_t}$ vector, which is ownership at thirty.

The assumptions on the structure of the measurement error are such that the limiting distribution of the variance-covariance matrix of measurement errors is

$$\begin{pmatrix} \sigma_v^2 & \mathbf{0}' \\ \mathbf{0} & \Sigma_{uu} \end{pmatrix}$$

with **0** in this matrix again following from having sampling errors associated with independent samples and Σ_{uu} being the covariance matrix of the measurement error in the *x* variables (in our case, two of the *x* variables, the time trend and the dummy for year 1981 or after, are based on sample year and so measured without error).

In this case, the OLS estimator based on observed variables is again inconsistent, with

$$\text{plim } \boldsymbol{\beta^{OLS}} = (\Sigma_{x^*x^*} + \Sigma_{uu})^{-1}\left(\Sigma_{x^*x^*}\boldsymbol{\beta} - \boldsymbol{\sigma_{uu_1}}\right)$$

and $\boldsymbol{\sigma_{uu_1}}$ being the first column of Σ_{uu}, that is, the variance and covariances between the measurement errors in x_1 and in each *x*-variable.

The feasible consistent estimator we use is then

$$(X'X - TS_{uu})^{-1}\left(X'(\boldsymbol{y} - \boldsymbol{x_1}) + T\boldsymbol{s_{uu_1}}\right)$$

with X the x-variables stacked up by year and T the number of years (28 in our data), S_{uu} a consistent estimate of Σ_{uu} and $\boldsymbol{s_{uu_1}}$ its first column. Again, these estimates come from the pooled micro-data.

The results based on the estimators just described are presented in columns (5) and (6) of Table 3, along with bootstrapped confidence intervals (based on 1000 replications). The coefficients on ownership at thirty are now -0.798 and -0.803, respectively, for the specifications without and with additional controls. These results are very much in line with the instrumental variables estimates. Both suggest a relatively small (exaggerating) net effect of biases due to measurement error on the OLS estimate.

[Table 3 about here]

Taken together, these estimates, using two different methods to correct for possible measurement error bias, suggest a very robust result. There is substantial catch up, and cohorts with low homeownership rates at thirty have closed about 80% of the "ownership-gap" by the time they reach age forty. Thus, during our sample period of 1970 to 2007, the observation of a low ownership rate at age 30 was not a strong indicator that ownership would also be low at 40. A caveat to this result is that description of the latest data by Hood and Joyce (2013), suggests that the cohort that reached age 40 in around 2012 had a low ownership rate at age 30 and at age 40; it remains to be seen whether this is a new trend or a feature of the unusual (Great Recession) period in which this cohort lived out its 30s.

To check the sensitivity of our results and assess the role of other determinants of catch-up, we include additional regressors in the baseline specification of Table 3. We choose regressors that are most in line with those used for ownership at thirty, bearing in mind that the catch-up regressions are done at the cohort level rather than at the individual level. We include, in turn and then all together, variables that are intended to describe the most favourable housing and credit market conditions that cohorts faced during their thirties. These

variables, with descriptive statistics in online Appendix Table A.3, are: the minimum ratio of house price to income between age 31 and age 40, where house prices are the English annual first-time buyer levels, and income is the annual average from 31 to 40 year olds; the minimum interest rate between age 31 and age 40; and, the maximum level of the credit conditions index (CCI) between age 31 and age 40.[23] Table 4 reports the results of the measurement error correction specification; we prefer this method for dealing with endogeneity because, while qualitatively similar results were found if we used two stage least squares, the coefficient on the excluded variable becomes insignificant in the first stage in the cases in which the credit conditions index is used as a control (columns (3) and (4) in Table 4).

[Table 4 about here]

In contrast to our findings regarding ownership at 30, it is the interest rate and credit conditions variables (rather than the house price variable) that are sometimes significant in the regressions reported in Table 4. On the other hand, the additional regressors do not account for our catch up results in the sense that they do weaken the coefficients on ownership at thirty. Indeed, adding a control for the minimum ratio of the house price to income between age 31 and 40, as a measure of affordability of housing, does not affect our baseline estimates of catch up, giving a coefficient of -0.833 on ownership at 30. With the addition of the minimum interest rate or the maximum CCI, or all additional controls together, the magnitude of point estimate for the coefficient on ownership at 30 actually increases, but we can never reject the hypotheses of full catch up or of equality with the results found in our baseline analyses. Thus our interpretation is that our catch-up findings are robust to adding these extra controls. The unusually long data set available given our adoption of synthetic cohort techniques, was crucial for finding these results. While it would be interesting to say more about the determinants of catch up, we feel that this would require further innovations of data or method.

V. EARLY OWNERSHIP AND PERSISTENT EFFECTS ON OTHER OUTCOMES

While the preceding analysis supports the hypothesis of full or near full "catch up" in homeownership rates, it does not rule out the possibility that ability to get on to the housing ladder by age thirty persistently affects other outcomes, such as the amount of housing assets that cohorts are ultimately able to purchase and/or the share of income spent on housing. To investigate this we continue to use the synthetic cohort data that we used for our catch up regressions, but we change the outcomes for the cohort at age 40 which we are relating to ownership rates at 30. In particular, we consider the average number of rooms (a proxy for house size) in the accommodation owned at age 40, the proportion of outright owners by age 40, and the ratio of the (average) mortgage payment to (average) income,[24] as dependent variables that we regress on the cohort ownership rate at age 30. The idea of these regressions is to consider whether earlier access to ownership is related subsequent advantages for those who own, and so the outcome variables are averages among 40 year-old home owners. Table 5 shows descriptive statistics for these outcome variables, while results from the regressions are shown in Table 6.

[Table 5 about here]

Table 6 shows results for the simplest specification with one regressor, but coefficients are not substantially affected if we add other regressors. Table 6 also reports results only for the "measurement error correction". Similar results are found if we use an instrumental variables strategy, but we prefer to report the measurement error correction both for consistency with Table 4, and because in the regression for number of rooms (where three observations are lost due to top-coding resulting in non-comparability in the data) the excluded variable (ownership at 29) is not significant in the first-stage.[25] It should be noted that since we are no longer considering a "catch-up" regression, the "measurement error correction" is now implemented by directly applying results from Deaton (1985). Thus, for

the univariate case that we report (and using the notation of Section IV above), the estimator[26] is:

$$\beta^{FLS} = \left(\sum (x^2 - s_u^2) \right)^{-1} \sum (yx)$$

[Table 6 about here]

The positive coefficient in the number of rooms regression suggests that home-owners from cohorts that access ownership earlier may end up owning slightly bigger houses at age 40. The coefficient of 4.767 suggests that a one standard deviation increase in home ownership at age 30 (an increase of 5.5 percentage points, see Table 5) is associated with approximately an extra quarter of a room (on average) in the home owned at age 40. Based on our results for outright ownership and for the mortgage payment to income ratio, we cannot reject the hypothesis that earlier access to ownership is not related to earlier outright ownership, or to mortgage payments being a lower drain on income by age 40.[27]

The pattern of coefficients across the three regressions is therefore consistent with the idea that those who access ownership early exploit any advantage to buy bigger properties (either straight away or through upsizing) rather than to reduce the burden of mortgage debt. However, we should be cautious drawing such a conclusion as a lack of significance in regressions could simply indicate that we do not have enough data to test hypotheses.

The results in this section necessarily relate to a limited set of (three) outcomes, the choice of which is largely data driven. Our analyses do not, for example, allow us to address whether those who access the housing market later end up buying in locations that have lower prices because they are less desirable in some dimension (perhaps in terms of local services or of commuting distances). Such interesting questions are left for further work.

To get a fuller picture of the "persistence" of relationships, we might also like to test hypotheses at later ages than 40. This could be particularly valuable for the "outright ownership" variable, since most households pay off their mortgages rather later than at age

40 (in our sample only around 8 to 9 per cent are outright owners by 40, see Table 5). To this end we tried our regressions for outcomes measured at age 50, but reducing the sample size by 10 observations compared to Table 6 resulted in noisy coefficient estimates. Our main "positive" result for these other outcomes is thus that we find some evidence that earlier access to home ownership may result in the ownership of slightly larger homes at around age 40.

VI. CONCLUSIONS

Due to the fact that England experiences significant house price volatility, with booms and busts, different birth cohorts have experienced very different housing market conditions in early adulthood. It is natural to ask whether these fluctuations have been associated with different homeownership outcomes for the birth cohorts that experienced them, and whether the differences, if present, persist into later life.

There are number of ways that one could address these questions. In this paper we have investigated these questions empirically, employing successive FES/EFS surveys over almost forty years, in conjunction with synthetic cohort methods. These data and methods allow us to track the ownership rates of different birth cohorts over a time period that captures three housing booms, and two housing busts.

We find that, over the past forty years, ownership rates at age thirty have varied substantially across birth cohorts. This variation is negatively correlated with house prices, but the relationship seems stronger before 1990 than subsequently. These patterns are common to men and women, and to the different regions of England. Overall, our results suggest that when a birth cohort faces a house price to income ratio that is one standard deviation above trend in early adulthood, then the homeownership rate of that birth cohort at age thirty is approximately 4 percentage points lower.

There has been a secular decline in ownership at age thirty from the early-1990s on. This is associated with a coincident decline in the fraction of thirty year olds in couple households. It could be that causality runs from household formation to housing demand, or from housing prices or supply to household formation, or both, or neither.

Those birth cohorts that were less likely to get onto the ladder by thirty were nonetheless subsequently able to "catch up", to a large degree, with cohorts that experienced more favourable initial conditions. Measurement error means that the raw correlation between ownership at thirty and subsequent growth in ownership, may misstate the true degree of catch up. Nevertheless, two different econometric methods which address that problem, and ancillary evidence, suggest that in our sample period (1970 – 2007) the apparent catch up is real. Cohorts with low homeownership rates at thirty closed about 80% of the "ownership-gap" by the time they reach age forty.

While there is full or near full "catch up" in homeownership rates, it is possible that ability to get on to the housing ladder by age thirty persistently affects other outcomes. We focus on housing outcomes. Our results are consistent with the idea that those who access ownership early exploit any advantage to buy bigger properties (either straight away or through upsizing) rather than to reduce the burden of mortgage debt. The "positive" result is that earlier access to home ownership may result in the ownership of slightly larger homes at around age 40.

As with any analysis, ours has limitations. An obvious limitation of the analysis in this paper is that it only documents the association of housing market conditions with the experiences of successive cohorts of young adults, and stops short of drawing causal inferences. While these associations are certainly suggestive of an effect of housing market conditions on outcomes, it is quite possible that there are important effects that run in the opposite direction - from the size and characteristics of different birth cohorts reaching young adulthood to housing market conditions. Disentangling these different effects is important,

but beyond the scope of this paper. It is nevertheless important to document, as this paper does, the key facts on cohort homeownership, both as a basis for further empirical work, and as targets for any structural modeling to match.

ACKNOWLEDGMENTS

This research has evolved from earlier work funded through the National Housing and Planning Advice Unit (NHPAU) Research Innovation Fund (Bottazzi et al., 2010). Bottazzi and Wakefield also acknowledge financial support from MIUR- FIRB 2008, project RBFR089QQC-003-J31J10000060001 and MIUR-PRIN 2010-11, project 2010T8XAXB_006. We also acknowledge co-funding form the ESRC-funded Centre for Microeconomic Analysis of Public Policy at the Institute for Fiscal Studies (grant number RES-544-28-5001.) We thank: Carl Emmerson, Cormac O'Dea, our editor (Alex Michaelides) and three anonymous referees; seminar participants at the University of Bologna, at the Royal Economic Society Annual Conference 2011, at the University of Milano-Bicocca, and at Ca' Foscari University of Venice, for helpful comments; and especially Anthony Murphy for comments and for providing the credit conditions data. All errors are our own.

REFERENCES

Aron, J., J. V. Duca, J. Muellbauer, K. Murata and A. Murphy (2012), "Credit, Housing Collateral, and Consumption: Evidence from Japan, The U.K., and The U.S", *Review of Income and Wealth*, 58(3), pp. 397–423.

Banks, J., R. Blundell, Z. Oldfield and J. Smith (2010), "House Price Volatility and the Housing Ladder", Working Papers 786, RAND Corporation.

Benito, A. (2009), "Who Withdraws Housing Equity and Why?", *Economica*, 76, pp.51–70.

Bottazzi, R. (2004), "Labour market participation and mortgage-related borrowing constraints", IFS Working Paper 04/09, Institute for Fiscal Studies, London.

Bottazzi, R., Crossley, T.F. and M. Wakefield (2010), "How do Housing Price Booms and Busts Affect Homeownership for Different Birth Cohorts?" National Housing and Planning Advice Unit Report.

Bottazzi, R., H. Low and M. Wakefield (2007), "Why do home owners work longer hours?", IFS Working Paper W07/10.

Bourassa, S. and M. Hoesli, (2010). "Why Do the Swiss Rent?", *The Journal of Real Estate Finance and Economics*, 40(3), pp. 286-309.

Cameron, G., J. Muellbauer, and A. Murphy, 2006, "Was there a British House Price Bubble? Evidence from a Regional Panel," CEPR Discussion Paper No. 5619, Centre for Economic Policy Research, London.

Chiuri, M.C. and T. Jappelli (2003), "Financial market imperfections and homeownership: A comparative study", *European Economic Review*, 47(5), pp. 857-875.

Chiuri, M.C. and T. Jappelli (2010), "Do the elderly reduce housing equity? An international comparison", *Journal of Population Economics*, 23(2), pp. 643-663.

Crossley, T.F. and Y. Ostrovsky (2003), 'A Synthetic Cohort Analysis of Canadian Housing Careers' SEDAP Research Paper #107, McMaster University.

Deaton, A. (1985), "Panel Data From a Time Series of Cross-Sections," *Journal of Econometrics*, 1985, pp. 109-126.

Dettling, L. and M.S. Kearney. (2014), "House Prices and Birth Rates: The Impact of the Real Estate Market on the Decision to Have a Baby," *Journal of Public Economics*, 110, pp.82-100.

Disney, R. and G. Luo, (2015), "The right to buy social housing in Britain: a welfare analysis", IFS Working Paper W15/05.

Duncan, G. and Hill, M., (1985). "Conceptions of Longitudinal Households: Fertile or Futile?" *Journal of Economic and Social Measurement*, 13, pp. 361-375.

Ermisch, J. (1999), "Prices, parents and young people's household formation," *Journal of Urban Economics*, 45, pp.47-71.

Fernandez-Corugedo, E. and J. Muellbauer (2006), *Consumer credit conditions in the United Kingdom*, Bank of England Working Paper 314, (http://www.bankofengland.co.uk/publications/workingpapers/wp314.pdf)

Hilber, C.A.L. and L., Yingchun, (2008.) "Explaining the black-white homeownership gap: The role of own wealth, parental externalities and locational preferences," *Journal of Housing Economics*, 17(2), pp. 152-174.

Hood, A. and R. Joyce, (2013), "The Economic Circumstances of Cohorts Born between the 1940s and the 1970s", IFS Report R89, Institute for Fiscal Studies: London.

Hurst, E. and F. Stafford, (2004). "Home Is Where the Equity Is: Mortgage Refinancing and Household Consumption," *Journal of Money, Credit and Banking*, 36(6), pp.985-1014.

Jappelli, T. (1999), "The Age-Wealth Profile and the Life-Cycle Hypothesis: A Cohort Analysis with Time Series of Cross-Sections of Italian Households," *Review of Income and Wealth*, 45(1), pp.57-75.

Kapteyn, A., R. Alessie and A. Lusardi (2005), "Explaining the wealth holdings of different cohorts: Productivity growth and Social Security", *European Economic Review*, 49, pp.1361 – 1391.

Linneman, P. and S. Wachter, (1989), "The Impacts of Borrowing Constraints on Homeownership," *Journal of the American Real Estate and Urban Economics Association*, 17(4), pp. 389-402.

Lusardi, A., and O.S. Mitchell, (2007), "Baby Boomer retirement security: The roles of planning, financial literacy, and housing wealth", *Journal of Monetary Economics*, 54(1), pp.205-224.

Myers, D., (1990). "Introduction: The Emerging Concept of Housing Demography," in Housing Demography: Linking Demographic Structure and Housing Markets (ed. D. Myers). The University of Wisconsin Press, Madison.

Myers, D., (1999), "Cohort Longitudinal Estimation of Housing Careers," *Housing Studies*, 14(4), pp.473-490.

Myers, D., (2001), "Reply to Frans Dieleman: Better Use of Cross-sectional Data," *Housing Studies*, 16(1), p.123.

NHPAU (2008), "Impact of worsening affordability on demand for social and affordable housing: tenure choice and household formation", NHPAU Research Findings, Number 2.

Appendix

a. Data details

(i) House-price data

Throughout this paper the house price data that we have used are based on house price indices that are now published by the Office for National Statistics (ONS), having transferred from the Department for Communities and Local Government (DCLG) from February 2012. The data are freely available via http://www.ons.gov.uk/ons/taxonomy/index.html?nscl=House+Price+Indices#tab-data-tables. The main indices that we use are published in ONS House Price Indices reference tables 8, 9 and 10 (previously DCLG live tables 590, 592 and 594). The indices are mix adjusted, where the mix adjustment is to allow for the fact that the composition of house types traded in the housing market changes from year to year (e.g. in some periods a higher proportion of large detached properties are traded, at other times flat sales are more important).

Since we investigate entry in to the housing market, in most of our analyses we use house price indices for properties bought by first-time buyers (ONS tables 9 and 10). These first time buyer indices are published from quarter 1 of 1983. In order to have a longer time series we "backcast" the first time buyer indices using growth rates from the all properties index for the relevant geographical area. Our main results are not substantially affected by avoiding this backcasting either by starting the analysis from 1983 or by using an all properties house price index (see also the discussion around Table 1 in Section III; full details on request).

With the published all properties price indices, and our extended first-time buyer indices, we have quarterly data beginning in quarter two of 1968. The data include separate series for the UK and for Great Britain, for England, Scotland, Northern Ireland and Wales, and for English regions (nine "government office regions"). We generally use the series for England, or the series for English regions, depending on which is appropriate to the analysis.

To convert from the price index into a price level, we use mix-adjusted prices for quarter 1 (February) of 2002, which are also published by and freely available from ONS (ONS table 13, previously DCLG live table 508).

In much of the paper, and many of our regression specifications, the house price variable that we use is a house price to average income ratio. To compute these ratios we divide the house price by average family income in our FES/EFS data. Generally we take average incomes at the region and quarter level (though some figures display statistics at national and/or annual level and in these cases average incomes are consistent with this). For the analyses of ownership at age 30 the price to income ratio is the relevant house price (the

regional first time buyer price in our main specifications) divided by the average family income of individuals around age 30; to get a reasonable cell size for taking average income by region and quarter we average across individuals aged 26 – 34 (inclusive). For our synthetic cohort regressions the object of analysis is changes after 30 and up to 40 (rather than levels at age 30), and so in those specifications where it is used the house price to income variable is defined in terms of the incomes of those aged 31 to 40, again averaged at the region and quarter level.

In places we use (the log of) real house prices. To convert nominal house prices in to real prices we deflate using the all-item Retail Prices Index to December 2007 levels (2007 is the latest year in our FES data on ownership). The RPI data are published by the Office for National Statistics, and we have monthly data. To deflate the quarterly house price series we use the (mean) average of the RPI for the 3 months corresponding to each quarter.

Though the house price data start in quarter two of 1968, not all the English regions have data for the full period due to changes in the drawing of regional boundaries. In particular the North-East, East (i.e. East Anglia) and South East series are available from quarter two of 1992, while the North West series is available from quarter one of 1999. When exploiting regional data, we either drop region-years in which the house price is not available, or, for the figures plotting the house price and ownership in broad English regions, we construct the price series based on only a subset of the more narrow regions that are the constituent parts of our broader regions. To ensure that changes in data availability by region do not generate changes in cell composition in our synthetic cohort results, in those analyses we always use national house price series that are available in all periods.

(ii) Credit Conditions Index

As mentioned in the main text, the Credit Conditions Index (CCI) that we use in specification 3 of table 1 and in specifications 3 and 4 of table 4, comes from Fernandez-Corugedo and Muellbauer (2006). Their index is constructed, for the period from the mid-1970s to the early 2000s, as the common underlying influence on ten measures of credit conditions. These measures include aggregate unsecured debt and mortgages, and age and region specific measures of the fraction of high loan-to-income, and value-to-income, mortgages. The ten equation system that is estimated includes controls for a comprehensive set of economic and demographic influences on the demand and supply of credit. Thus the unified CCI that is derived captures the common variation in the ten credit indicators which cannot be explained by the economic and demographic controls. It is this index that we use in

our analysis. Full details of it and its construction are provided in Fernandez-Corugedo and Muellbauer (2006).

b. Synthetic-Cohort Analysis: checking for group consistency

As mentioned above, before we apply synthetic cohort analyses to regional samples, we need to conduct some checks on the data to make sure it is valid to do so.

The cohort methods hinge on cohort composition remaining constant over time. Random samples of fifty year-olds in 1980 and 60-years olds in 1990 are informative about the average experience of individuals in the 1930 birth cohort if the set of people in the population with that characteristic (born in 1930) is fairly constant between 1980 and 1990. If that is not the case then changes in the homeownership rate between 1980 and 1990 will confound changes in the homeownership rate among the individuals that the 1980 sample was drawn from with changes in the composition of the cohort.

At a national level, the main threats to the validity of this assumption are (i) immigration, (ii) emigration, and (iii) differential mortality. For example, suppose that the 1930 birth-cohort experiences some mortality between 1980 and 1990 and that this mortality is concentrated amongst those with lower socioeconomic status and wealth. As these people are less likely to own homes this can lead to a rise in the homeownership rate of the cohort even though there is no change in the homeownership probability of any given individual in the cohort. As we are ultimately interested in the life-course experience of individuals, we would consider this a spurious selection (or compositional) effect. Similar effects arise if, for example, immigrants who join a cohort as it ages have lower (or higher) homeownership rates than the native born.

Turning to regional analysis, we face two main difficulties. First, if we look at smaller regions then the available sample for any given birth cohort in any given survey year can be quite small. These small cell sizes then lead to considerable sampling variation in the homeownership rate of a given birth-cohort, at a given age, in a given region. The resulting age paths of homeownership are therefore potentially quite noisy, with meaningless year-on-year variations.

The second problem is that threats to the validity of the constant birth cohort composition assumption are potentially more severe at the regional level. This is because inter-regional migration might be greater than international migration.

There is a way to check these issues internally in the data. The idea is to use the data to track across age a characteristic (or characteristics) of a birth-cohort (or birth/region cohort) that we believe should be constant. If cohort composition does change over time, we might

expect this to be manifest in these age profiles. To implement this idea, we organized the data for England into three large regions (South, Midlands and North) and within each region, into 10-year birth cohorts. We then examine two features of each cohort as it ages: cohort size, and the fraction of individuals in the cohort who left full-time education at or after age 18. Changes in estimated cohort size would reflect mortality as well as migration into and out of the region. The fraction of individuals in the cohort who left full-time education at or after age 18 should of course be roughly constant after age 18 and if it changes as the cohort ages this would indicate either differential mortality or that higher (or lower) education individuals are being added (or subtracted) from the cohort by migration.

The results of this analysis are presented in Figures A.1 and A.2. As always we have been careful to use the survey weights in all calculations. Figure A.1 shows the estimated cohort size for a set of cohorts defined by region of residence (South, Midlands, North) and 10 year birth cohort (1930s, 1940s, 1950s, 1960s). Size is on the vertical axis and age on the horizontal axis. Vertical differences between cohort lines indicate "cohort effects." For example, particularly in the South, the 1960s birth cohort (in solid dark grey) is significantly larger than the 1930s cohort (in dashed light grey). The line for each cohort traces out the age profile for that cohort. This figure shows some mild decline in estimated cohort size as each cohort ages, which probably reflects a combination of mortality and net emigration. There is some suggestion of an accelerated decline in cohort size past age 65 (which we see in our data only for the 1930s cohort) which might be consistent with accelerating mortality or emigration associated with retirement.

[Figure A.1 and A.2 about here]

The key point that we draw from Figure A.1, however, is that changes in estimated cohort size are quite modest (at least before age 65) and very similar across birth cohorts and regions. We would have been rather more concerned if Figure A.1 showed cohorts in one region growing while cohorts in other regions shrank, indicating substantial net migration between regions. This does not, however, appear to be the case.

Figure A.2 follows the same pattern but traces out the fraction of individuals in the cohort who left full-time education at or after age 18. The age profiles of the different cohorts are fairly noisy, and perhaps exhibit some small upward trend with age. The latter would be consistent with differential mortality (higher socioeconomic status individuals having greater life-expectancy) and/or some incidence of older individuals returning to school. The main point again is that the age effects do not appear to be dramatic, and do not appear to differ significantly across regions.

It would certainly be possible to push this analysis further, for example by subjecting the age profiles apparent in these figures to formal statistical tests. But our conclusion from these figures is that analysis at the level of broad regions is feasible, and the constant composition assumption is no more dangerous at this level of region than at the level of England as a whole. On the other hand, the sampling variability in age profiles apparent especially in Figure A.2 suggests to us that, due to small sample sizes, analysis at the level of more disaggregated regions would not be advisable.

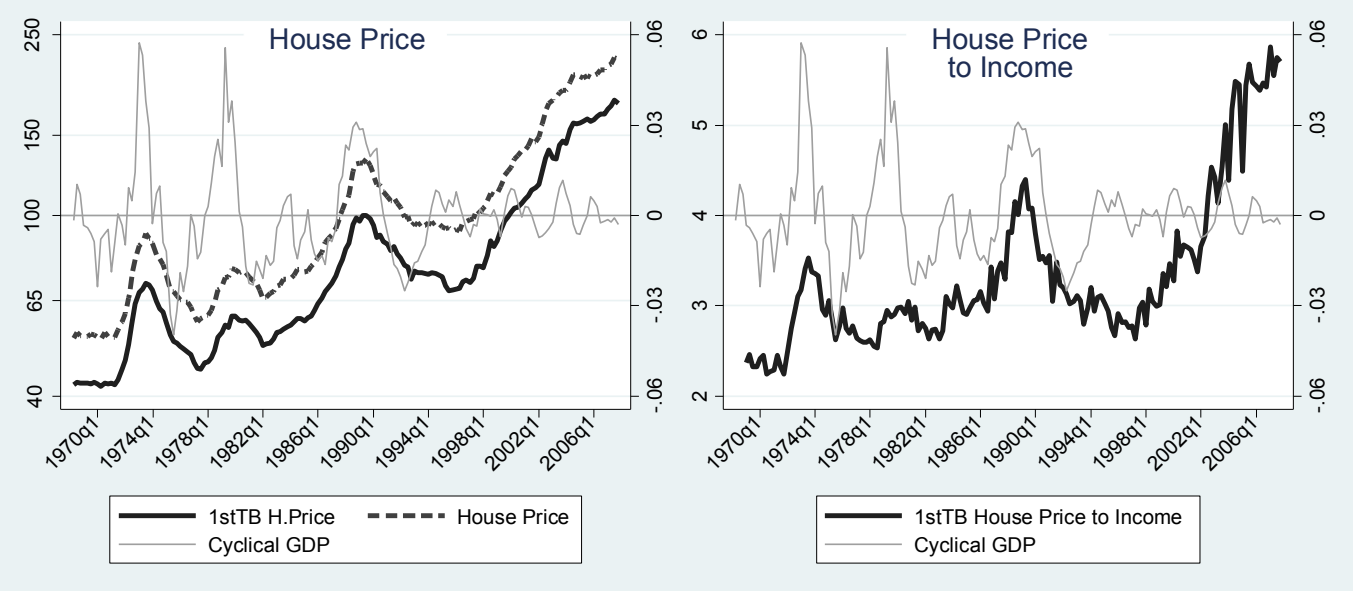
c. Further Descriptive Statistics

This subsection provides further detail on the pseudo-panel data constructed from successive FES/EFS surveys.

[Tables A.1 and A.2 about here]

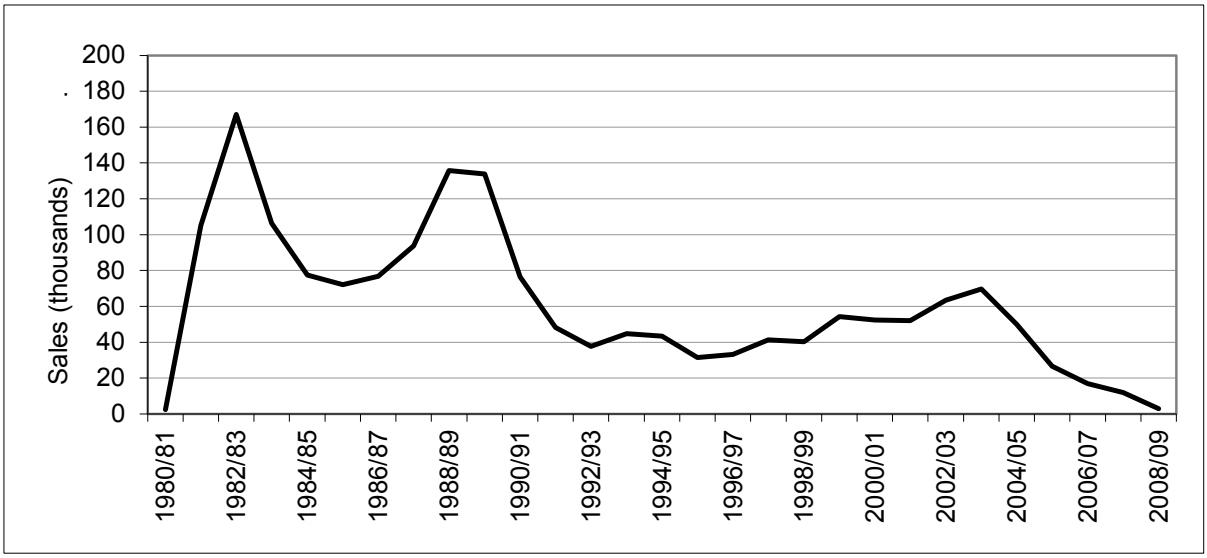
[Table A.3 about here]

Figure 1: Log Real House Prices, and the House Price to Income Ratio (for the "young"), with the Cyclical Component of GDP, England 1968 – 2007 (quarterly)



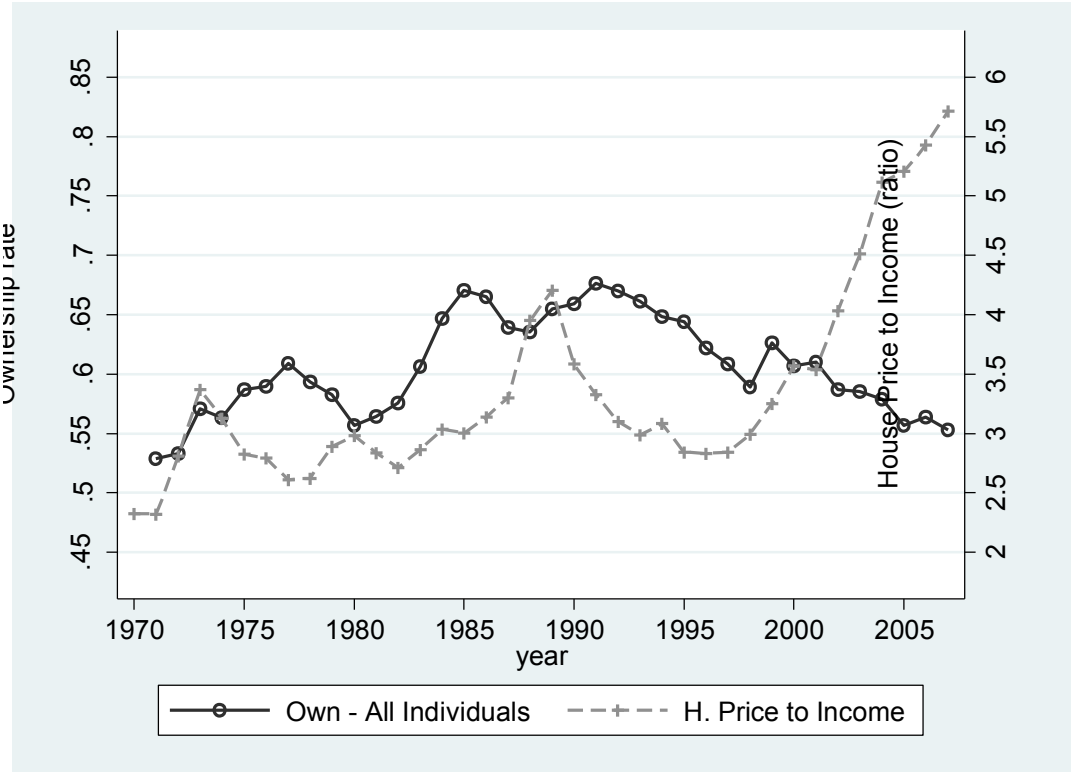
Notes: In the "House Price" panel the vertical axes are: real house prices in thousands of 2007 pounds (left-hand axis), and the cyclical component of log real GDP from the Hodrick-Prescott filter (right-hand axis). The real house price axis is plotted on a logarithmic scale but labeled in (2007) pounds. House prices are deflated by the authors using the all items retail prices index. In the "House Price to Income" panel the vertical axes are: the house price to income ratio (left-hand axis), and the cyclical component of log real GDP from the Hodrick-Prescott filter (right-hand axis). The house price to income ratio is the ratio of the first time buyer house price to average incomes around age 30 (ages 26 – 34) in our data, the same ratio that we use in the regressions reported in table 1. Fuller details of the data, and particularly of the house price data, are as described in the online technical appendix to the paper.

Figure 2: Local authority housing stock sold through the right to buy scheme in England, 1980 – 2009



Notes: This chart uses data and reproduces a figure that are published by the Department for Communities and Local Government: see chart 671 via http://www.communities.gov.uk/housing/housingresearch/housingstatistics/housingstatisticsby/socialhousingsales/livetables/ .

Figure 3: Proportion of individuals aged thirty who are owner-occupiers, and ratio of (first time buyer) house price to average income around 30, 1971-2007, England.



Notes: As in Figure 1 and in the regressions in Table 1, the house price to average income series is the ratio of the first time buyer house price to average incomes around age 30 (ages 26 – 34) in our data.

Figure 4: Proportion of individuals aged thirty who are owner-occupiers, with the same by couple status, and the proportion of thirty-year olds who are in couples, 1971-2007, England.

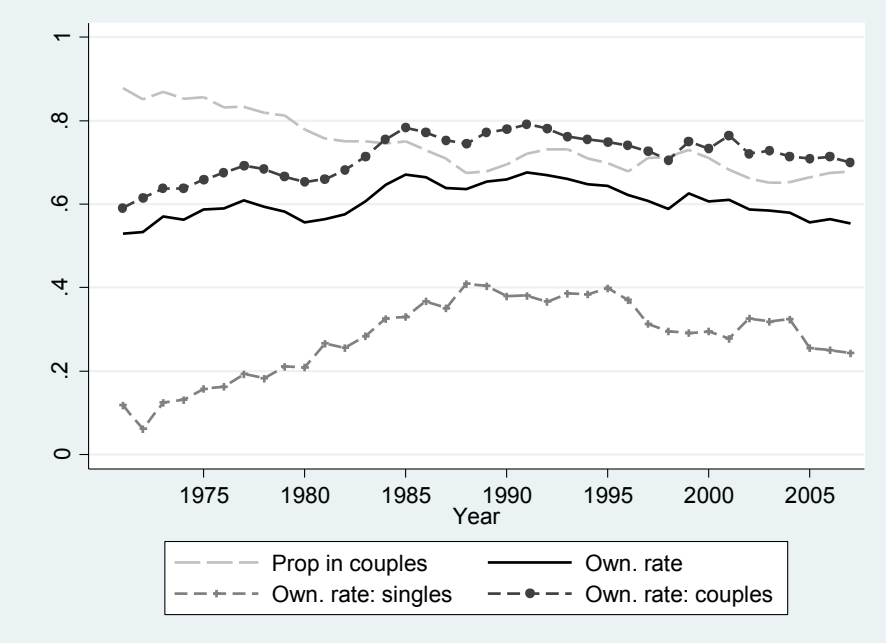


Figure 5: Ownership rates at different ages by year: all individuals, England

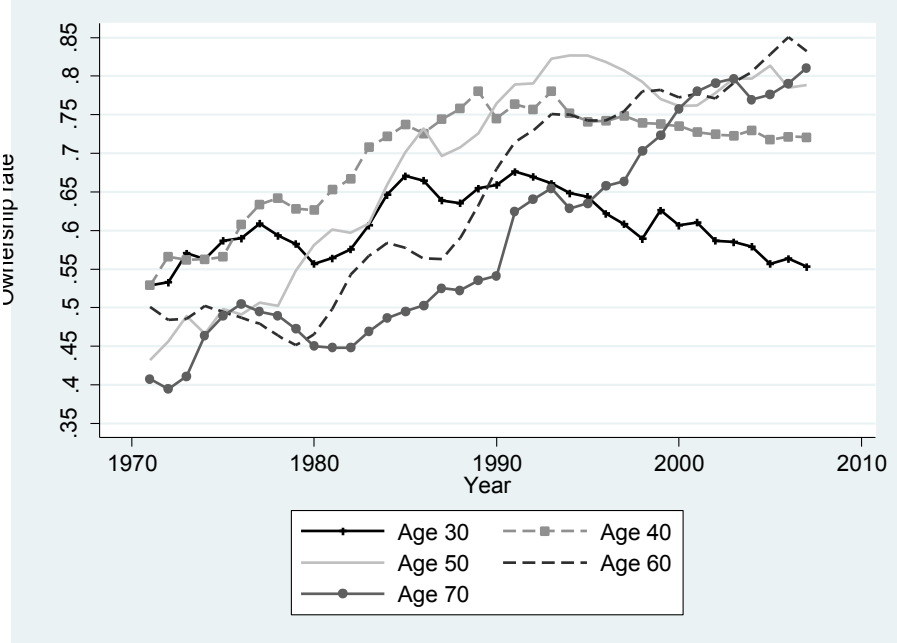


Figure 6: Catch up by forty?
Ownership change (rate at forty – rate at thirty), against ownership at thirty, England and regions

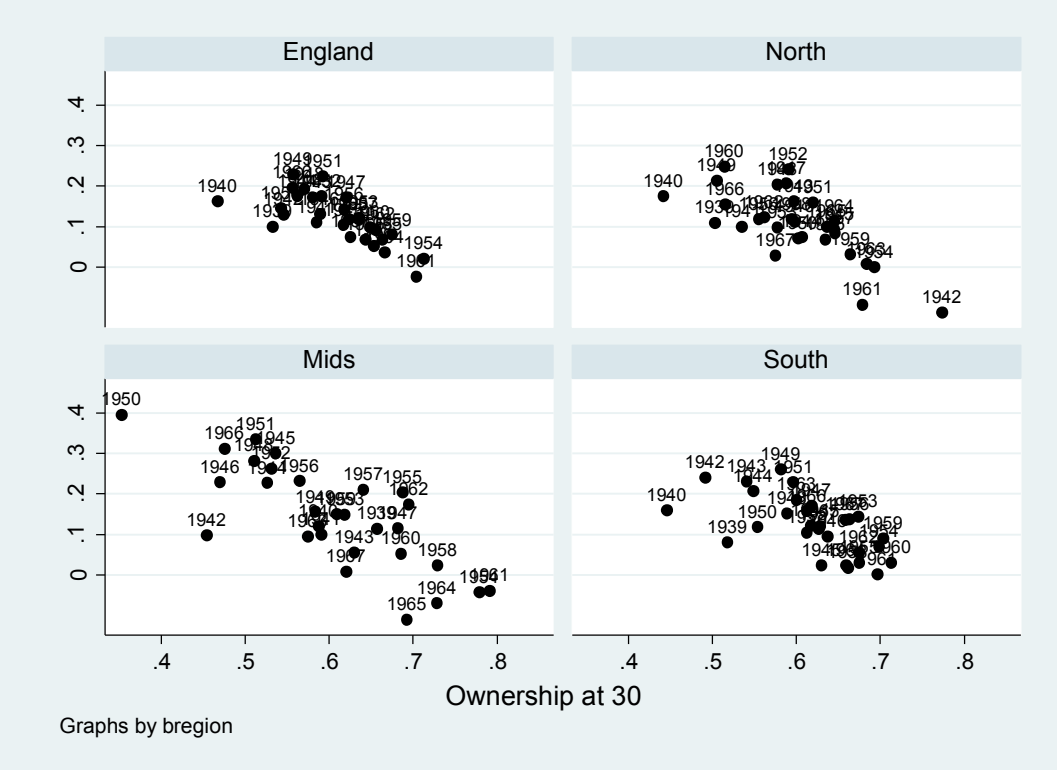


Figure 7: Catch up by fifty?
Ownership change (rate at fifty – rate at thirty), against ownership at thirty, England and regions

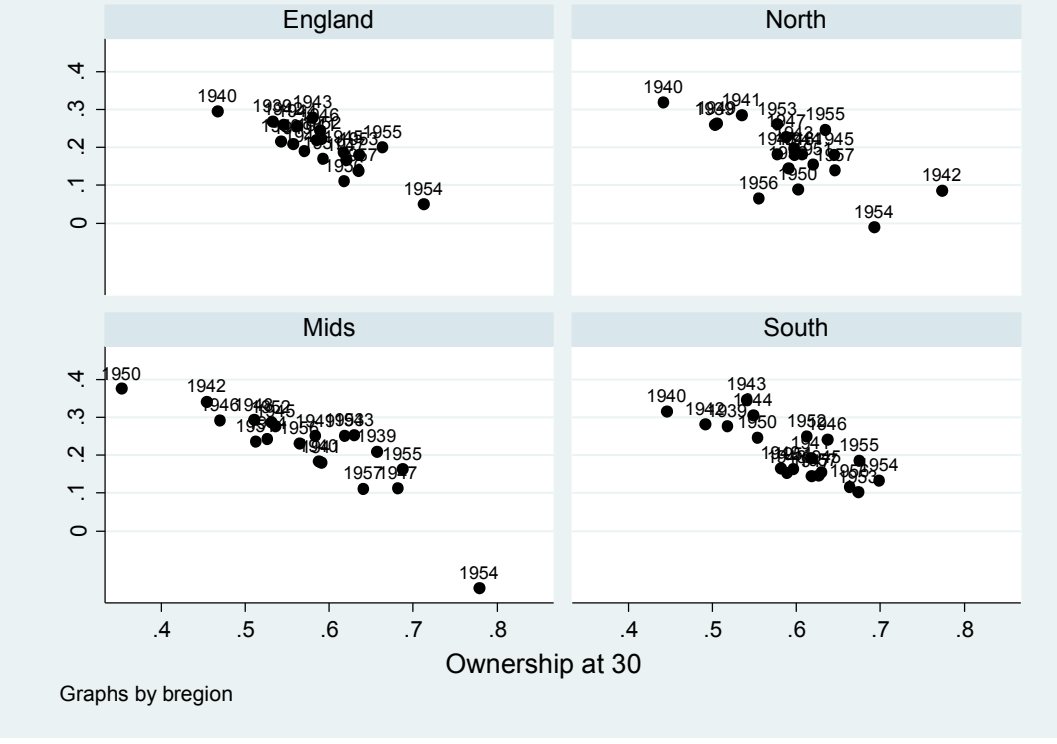


Figure A.1: Population sizes by broad region for (10year) cohorts

Figure A.2: Proportion who left full-time education at or after age 18

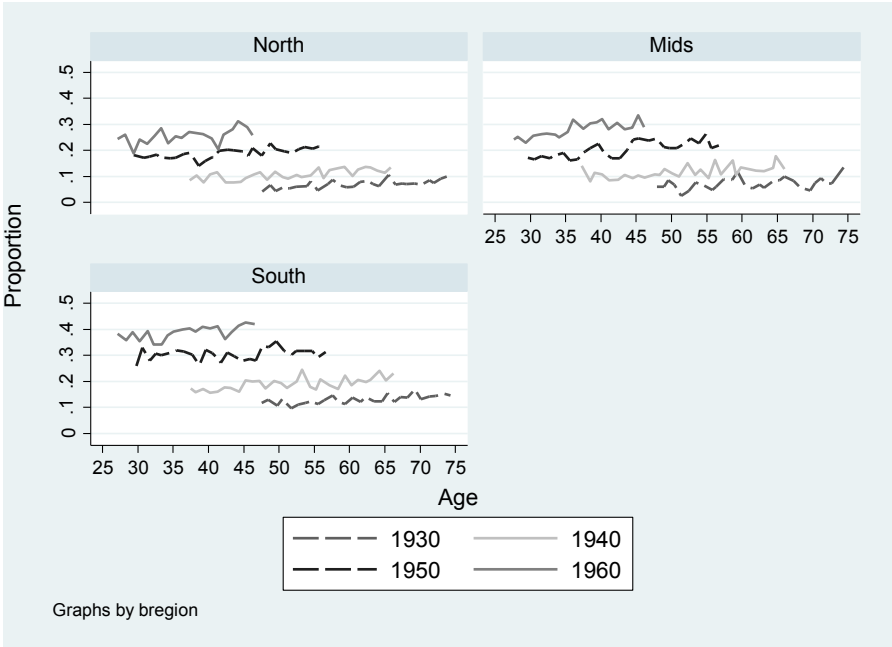


Table 1: Probit Regression for Ownership at Thirty, Pooled Data for England: Dependent Variable: Ownership at thirty

Regressor	Coeff (*s.e.*)	Marginal effect	Coeff (*s.e.*)	Marginal effect	Coeff (*s.e.*)	Marginal effect	Coeff (*s.e.*)	Marginal effect
House price / income	-0.1076 *** (*0.0264*)	-0.0414	-0.1140 *** (*0.0280*)	-0.0435	-0.1050 *** (*0.0367*)	-0.0400		
Nominal interest rate	0.0020 (*0.0095*)	0.0007	0.0014 (*0.0107*)	0.0005	0.0011 (*0.0010*)	0.0004	-0.0035 (*0.0113*)	-0.0013
Credit Conditions Index					-0.0855 (*0.4897*)	-0.0325	0.0979 (*0.5228*)	0.0373
RTB sales this year	0.2896 *** (*0.0585*)	0.1115	0.2575 *** (*0.0671*)	0.0983	0.2817 (*0.1801*)	0.1072	0.1848 (*0.1798*)	0.0704
Cumulative RTB sales	-0.0068 (*0.0069*)	-0.0026	-0.0088 (*0.0082*)	-0.0034	-0.0046 (*0.0256*)	-0.0017	-0.0256 (*0.0263*)	-0.0098
Log family income	0.9068 *** (*0.0676*)	0.3491	0.8955 *** (*0.0716*)	0.3417	0.8680 *** (*0.0720*)	0.3302	0.8720 *** (*0.0717*)	0.3319
Female (0/1)	0.1008 *** (*0.0311*)	0.0388	0.0656 * (*0.0338*)	0.0250	0.0686 * (*0.0362*)	0.0261	0.0690 * (*0.0368*)	0.0263
Couple (0/1)	0.6694 *** (*0.0560*)	0.2604	0.6305 *** (*0.0562*)	0.2439	0.6576 *** (*0.0555*)	0.2540	0.6569 *** (*0.0545*)	0.2538
Number of kids	-0.1295 *** (*0.0187*)	-0.0499	-0.1050 *** (*0.0228*)	-0.0400	-0.1169 *** (*0.0227*)	-0.0445	-0.1192 *** (*0.0226*)	-0.0454
Post-compuls Educ (0/1)			0.1968 *** (*0.0371*)	0.0751	0.1980 *** (*0.0393*)	0.0753	0.1843 *** (*0.0408*)	0.0701
Sample and sample size	1969 – 2007, 5687		1978 – 2007, 4623		1978 – 2005, 4328		1978 – 2005, 4328	
Pseudo r-squared	0.2103		0.2244		0.2190		0.2165	

Notes: (a) Eight region dummies and a constant are included in all specifications; **(b)** Standard errors clustered at the year level;
(c) Marginal effects measured at means of independent variables; **(d)** ***, ** and * respectively indicate significance at 1%, 5% and 10% levels.

Table 2 Descriptive Statistics for Ownership Rate, Selected Ages and Cohorts, England

	Mean	Standard Deviation	Median	Minimum observed	Maximum observed
(a) Ownership Rate at Age Thirty and Forty, Among Cohorts Observed at Both Ages **Note: Based on 29 observations at each age, (1979-2007 for age 40, 1969-1997 for age 30).**					
Age 40	**0.729**	**0.043**	**0.737**	**0.631**	**0.817**
Age 30	**0.609**	**0.055**	**0.618**	**0.468**	**0.713**
(b) Ownership Rate at Age Thirty and Fifty Among Cohorts Observed at Both Ages **Note: Based on 19 observations at each age, (1989-2007 for age 50, 1969-1987 for age 30).**					
Age 50	**0.793**	**0.036**	**0.801**	**0.729**	**0.863**
Age 30	**0.591**	**0.054**	**0.589**	**0.468**	**0.713**

Table 3: Catch up Regression Estimates
Dependent variable: change in ownership rate, age forty minus age thirty, England

	OLS		Two-Stage Least Squares		Measurement Error correction	
	(1)	(2)	(3)	(4)	(5)	(6)
	Coeff	Coeff	Coeff	Coeff	Coeff	Coeff
Regressor	(95% c.i.)	(95% c.i.)	(95% c.i.)	(95% c.i.)	(95% c.i.)	(95% c.i.)
Ownership at age 30	-0.871 ***	-0.951 ***	-0.835 **	-0.820 *	-0.798 ***	-0.803 **
	(-1.162,-0.579)	(-1.375,-0.527)	(-1.499,-0.172)	(-1.688,0.047)	(-1.206,-0.506)	(-1.675,-0.091)
Year 1981 or after (0/1)		0.019		0.006		0.021
		(-0.052,0.091)		(-0.094,0.107)		(-0.076,0.114)
Couple (0/1) at age 30		0.073		-0.042		0.190
		(-0.369,0.516)		(-0.392,0.476)		(-0.541,0.752)
Constant	0.654 ***	0.635 ***	0.632 ***	0.587 **		
	(0.475,0.833)	(0.245,1.025)	(0.226,1.039)	(0.136,1.037)		
Observations	28	28	28	28	28	28
R-squared	0.59	0.60				

Notes: (a) ***, ** and * respectively indicate significance at 1%, 5% and 10% levels.

(b) For columns (3) and (4) (two-stage least squares): coefficients of the second stage are reported. In the first stage ownership at thirty is regressed on the regressors included in the second stage, plus ownership at twenty-nine. Ownership at twenty-nine is the "excluded variable" and has a coefficient of 0.371 (s.e. 0.164) for the specification of column (3) and of 0.317 (s.e. 0.135) for column (4).

(c) For columns (5) and (6) (measurement error correction): **(i)** Details on the methodology adopted for the measurement error correction are provided in Section IV; **(ii)** data are in deviations from means and so there is no constant; **(iii)** reported confidence intervals and significance tests (stars on coefficients) are obtained by bootstrapping (1000 replications, strata: age and year of birth).

Table 4: Catch up Regression Estimates: Sensitivity to adding regressors
Dependent variable: change in ownership rate, age forty minus age thirty, England
Estimation method: measurement error correction

	(1) Coeff (95% c.i.)	(2) Coeff (95% c.i.)	(3) Coeff (95% c.i.)	(4) Coeff (95% c.i.)
M.E. Correction				
Ownership at age 30	-0.833 ** (-1.562, -0.247)	-1.174 ** (-1.849, -0.261)	-1.819 ** (-2.234, -0.267)	-1.414 ** (-2.011, -0.087)
Year 1981 or after (0/1)	0.019 (-0.079, 0.117)	0.031 (-0.093, 0.102)	0.107 (-0.070, 0.150)	0.096 (-0.031, 0.161)
Couple (0/1) at age 30	0.165 (-0.565, 0.767)	0.854 (-0.383, 1.333)	1.033 (-0.395, 1.479)	0.632 (-0.477, 1.302)
Min HP/income 31 – 40	0.016 (-0.061, 0.064)	. .	. .	0.037 (0.039, 0.070)
Min int. rate 31 – 40	. .	0.010 ** (0.001, 0.147)	. .	0.012 ** (0.002, 0.026)
Max CCI 31 – 40	. .	. .	0.271 * (-0.000, 0.354)	0.157 (-0.090, 0.300)
Observations	28	28	28	28

Notes: (a) ***, ** and * respectively indicate significance at 1%, 5% and 10% levels.
(b) Details on the methodology for the measurement error correction are provided in Section IV.
(c) Data are in deviations from means and so there is no constant.
(d) Reported confidence intervals and significance tests (stars on coefficients) are obtained by bootstrapping (1000 replications, strata: age and year of birth).

Table 5: Descriptive Statistics for Number of Rooms, Outright Ownership, and Mortgage Payment to Income
The statistics are for the distribution of averages for 40 year-old home-owners in our data for England

Variable	Mean	Standard Deviation	Median	Minimum observed	Maximum observed
Number of Rooms	6.092	0.264	6.002	5.689	6.539
Outright Ownership Rate	0.086	0.021	0.084	0.047	0.136
Average Mortgage Payment / Average Income	0.142	0.023	0.141	0.109	0.214

Notes: In line with the regressions in Table 6, for "number of rooms" there are 25 annual observations (1980 – 2004, the variable is top-coded from 2005); for the other variables there are 28 observations (1980 – 2007).

Table 6: Cohort Analyses of House Size, Outright Ownership and Mortgage Payment to Income
Dependent variables for owners at 40, England,
Estimation method: measurement error correction

Dependent variable:	Number of Rooms	Outright Ownership	Mortgg. Payment / Income (ratio)
	Coeff (95% c.i.)	**Coeff (95% c.i.)**	**Coeff (95% c.i.)**
M.E. Correction			
Ownership at age 30	4.767 *** (1.648, 5.532)	-0.079 (-0.267, 0.168)	-0.004 (-0.157, 0.142)
Observations	25	28	28

Notes: (a) ***, ** and * respectively indicate significance at 1%, 5% and 10% levels.
(b) Details on the methodology for the measurement error correction are provided in Section IV.
(c) Data are in deviations from means and so there is no constant.
(d) Reported confidence intervals and significance tests (stars on coefficients) are obtained by bootstrapping (1000 replications, strata: age and year of birth).
(e) All dependent variables are averages across home-owners at age 40.
(f) For the "number of rooms" regression, the 25 observations are from 1970 – 2004 (included): from 2005 the dependent variable is top coded and so no longer comparable to earlier waves.
(g) The mortgage payment to income variable is the ratio of average mortgage payment in the sample of 40 year-old owners, to average income in this sample. Averaging before taking the ratio improves robustness to the presence of "outlier" individual families with very high mortgage payment to income ratios (often due to low reported income).

Table A.1 Ownership Rates, Ages thirty-fifty

	Year of Birth																														
Year	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967
1969	.533																												
1970	.595	.468																											
1971	.512	.557	.586																										
1972	.578	.589	.602	.545																									
1973	.559	.598	.574	.58	.581																								
1974	.601	.606	.561	.664	.6	.563																							
1975	.635	.613	.653	.599	.574	.601	.617																						
1976	.629	.603	.632	.66	.642	.642	.585	.589																					
1977	.628	.644	.546	.613	.573	.69	.636	.639	.62																				
1978	.638	.621	.668	.642	.633	.624	.652	.679	.622	.571																			
1979	.632	.69	.672	.659	.738	.654	.633	.727	.588	.644	.557																		
1980	.683	.631	.701	.635	.69	.706	.681	.648	.656	.61	.67	.543																	
1981	.553	.636	.696	.667	.66	.659	.693	.651	.689	.639	.655	.62	.593																
1982	.668	.642	.68	.675	.66	.737	.72	.681	.693	.672	.681	.663	.696	.591															
1983	.687	.662	.738	.718	.753	.79	.721	.772	.778	.735	.663	.673	.739	.641	.635														
1984	.774	.719	.739	.717	.684	.737	.719	.698	.759	.693	.731	.71	.686	.61	.599	.713													
1985	.719	.766	.72	.693	.683	.741	.721	.755	.784	.736	.704	.762	.743	.721	.677	.613	.664												
1986	.73	.746	.75	.797	.79	.783	.771	.72	.684	.717	.746	.728	.761	.754	.702	.691	.669	.618											
1987	.786	.711	.766	.767	.78	.758	.817	.769	.793	.746	.739	.707	.742	.724	.737	.64	.683	.682	.635										
1988	.736	.709	.794	.737	.756	.81	.771	.757	.78	.762	.801	.717	.753	.707	.73	.636	.74	.702	.605	.653									
1989	.801	.697	.766	.76	.774	.802	.77	.792	.778	.758	.786	.789	.722	.78	.748	.724	.747	.703	.656	.699	.675								
1990		.762	.786	.786	.803	.836	.795	.796	.822	.814	.842	.688	.774	.78	.772	.749	.742	.739	.795	.702	.717	.649							
1991			.805	.762	.797	.783	.817	.798	.848	.726	.784	.767	.817	.829	.79	.765	.755	.736	.712	.724	.676	.743	.704						
1992				.805	.837	.762	.811	.846	.807	.798	.805	.761	.758	.766	.701	.699	.769	.679	.692	.767	.678	.676	.637	.656					
1993					.858	.796	.853	.824	.843	.769	.752	.794	.777	.75	.757	.686	.767	.721	.757	.674	.756	.684	.667	.671	.623				
1994						.818	.851	.833	.825	.803	.814	.83	.737	.826	.747	.734	.759	.817	.708	.696	.699	.75	.682	.684	.633	.666			
1995							.804	.775	.769	.787	.767	.867	.796	.724	.725	.769	.732	.772	.717	.682	.75	.73	.745	.679	.697	.616	.643		
1996								.833	.804	.834	.8	.733	.797	.803	.722	.752	.774	.761	.824	.679	.685	.675	.674	.69	.59	.608	.644	.556	
1997									.786	.818	.823	.793	.754	.778	.82	.779	.716	.744	.753	.729	.764	.754	.737	.695	.67	.648	.744	.585	.625
1998										.76	.802	.816	.801	.8	.774	.78	.758	.759	.76	.705	.667	.718	.725	.705	.699	.631	.711	.668	.629
1999											.766	.809	.78	.8	.825	.808	.8	.767	.756	.694	.755	.772	.752	.723	.768	.683	.649	.675	.668
2000												.758	.806	.783	.833	.747	.847	.803	.714	.743	.772	.746	.705	.756	.691	.712	.665	.744	.714
2001													.763	.772	.806	.786	.797	.823	.78	.779	.717	.694	.681	.764	.714	.717	.669	.705	.697
2002														.813	.843	.822	.844	.794	.829	.776	.795	.736	.753	.747	.753	.745	.702	.736	.632
2003															.815	.828	.814	.784	.776	.802	.792	.769	.777	.769	.741	.737	.719	.782	.76
2004																.763	.787	.771	.829	.794	.743	.765	.716	.759	.73	.701	.686	.71	.69
2005																	.863	.75	.79	.813	.791	.766	.728	.737	.717	.749	.712	.697	.763
2006																		.729	.797	.815	.794	.779	.749	.755	.775	.782	.75	.752	.731
2007																			.772	.793	.736	.758	.781	.766	.72	.716	.737	.725	.7

Table A.2 Cell Sizes

year	Year of Birth																														
	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967
1969	222																												
1970	183	180																											
1971	239	222	222																										
1972	187	214	215	221																									
1973	204	189	212	202	245																								
1974	191	187	192	214	229	210																							
1975	208	208	196	206	247	232	247																						
1976	204	201	185	219	188	233	241	235																					
1977	199	210	184	210	204	204	244	272	325																				
1978	178	184	184	193	228	197	202	203	301	282																			
1979	194	162	185	188	204	224	217	216	251	266	256																		
1980	171	171	170	217	219	250	239	250	262	269	207	234																	
1981	209	192	184	219	221	248	274	247	322	302	273	273	247																
1982	191	212	190	199	215	218	267	261	334	280	252	252	255	220															
1983	176	186	200	168	206	212	228	219	309	262	261	234	229	223	227														
1984	174	177	189	190	204	218	198	175	230	230	227	240	239	243	199	246													
1985	164	166	158	190	191	165	196	215	275	262	259	235	238	217	215	215	228												
1986	151	155	166	190	191	218	182	235	234	223	219	223	228	211	226	229	236	234											
1987	157	167	141	164	199	202	199	208	247	267	234	241	236	219	254	210	235	229	259										
1988	169	170	150	183	167	188	226	240	265	256	227	218	217	197	249	242	202	216	212	211									
1989	162	161	157	163	182	197	220	222	269	237	250	205	230	186	216	244	204	212	228	216	230								
1990		136	135	144	172	204	176	179	230	226	202	211	188	199	186	193	228	222	218	202	246	223							
1991			134	155	181	178	200	208	202	179	200	178	202	204	211	199	198	199	234	234	209	208	196						
1992				155	211	179	203	200	273	229	195	218	189	202	220	214	211	220	246	235	248	238	267	243					
1993					167	161	176	176	234	222	190	175	192	163	198	194	193	212	212	216	240	196	245	216	265				
1994						170	149	164	229	220	203	182	192	182	178	197	195	198	211	227	216	207	198	225	222	228			
1995							144	188	219	213	181	182	188	165	198	202	178	187	205	210	209	212	229	225	236	255	235		
1996								172	196	193	189	149	160	173	153	190	180	187	184	202	235	199	213	251	238	241	241	224	
1997									204	168	160	180	162	170	152	181	180	176	185	178	228	211	234	229	222	202	241	211	194
1998										159	170	145	169	153	146	152	152	199	175	194	172	189	197	186	209	212	212	199	194
1999											165	158	178	131	169	180	166	166	177	179	197	185	177	202	204	226	207	227	246
2000												179	160	152	163	180	165	152	152	155	151	164	195	163	189	181	218	217	203
2001													174	164	178	180	174	206	185	203	196	210	249	204	237	237	242	244	205
2002														177	181	157	155	176	170	191	160	192	221	197	210	216	210	214	228
2003															156	174	159	172	200	169	188	189	193	206	227	246	223	207	190
2004																163	153	177	175	186	186	190	195	216	213	224	195	191	190
2005																	141	155	166	180	183	191	188	178	212	198	217	212	217
2006																		133	169	178	189	162	189	200	210	186	192	202	169
2007																			165	149	168	158	163	190	169	161	164	175	153

Table A.3: Descriptive Statistics for Regressors in the Cohort Regressions

Variable	Mean	Standard Deviation	Median	Minimum observed	Maximum observed
Couple (0/1) at age 30	0.761	0.067	0.751	0.659	0.893
Min HP/income 31 – 40	2.594	0.107	2.589	2.326	2.869
Min int. rate 31 – 40 (percent rate)	6.050	1.975	5.201	3.544	9.263
Max CCI 31 – 40	0.811	0.212	0.871	0.120	1

Notes: There are 28 annual observations, in line with the sample from the regressions reported in Tables 3 and 4 and in the final two specifications in Table 6. For the final independent variable, 61% of the sample is age 30 in or after 1981.

---

[*] Author order is alphabetical. Previous versions of this work were circulated under the titles "House Prices and Homeownership: a Cohort Analysis" and "Late starters or excluded generations? A cohort analysis of catch-up in home ownership in England".

[1] Sources: Department for Communities and Local Government (house prices) and ONS (Average Earnings).

[2] There is concern with "generation gaps" in home ownership. See for example, http://www.ft.com/cms/s/0/52a03274-bce1-11e4-a917-00144feab7de.html.

[3] The name of this survey changed in 2001, although the content and design largely continued, so that it is possible to construct a consistent series.

[4] Where a family is adult singles or couples, plus any dependent children. In the UK this corresponds to a benefit unit.

[5] We in fact use their credit conditions index in Section III.

[6] Some local schemes had existed in the 1970s.

[7] For the raw data, see chart 671 via http://www.communities.gov.uk/housing/housingresearch/housingstatistics/housingstatisticsby/socialhousingsales/livetables/

[8] This number of sales is a product of both the policy, and of the decision of tenants of whether to exercise their "right to buy". A full discussion of the policy (and this aspect in particular), is provided by Disney and Luo (2015).

[9] See: http://www.ons.gov.uk/ons/taxonomy/index.html?nscl=House+Price+Indices#tab-data-tables

[10] This measure was chosen because consistent data are available for a long time period. For the period in which data overlap it is very close to the Bank of England base rate.

[11] Full details are available on request.

[12] These dissagregations are not shown, but figures with these ownership rates by group are available in Bottazzi, Crossley and Wakefield (2010).

[13] As a sensitivity test, we have also estimated all three models on a common time frame. The results are similar and are available on request.

[14] We experimented with also including a variable indicating that an observation is from 1981 or after, but when included alongside the right to buy variables, all three were not separately statistically significant. We chose to keep the more interpretable measures of the impact of the policy.

[15] The credit conditions index used by Fernandez-Corugedo and Muellbauer (2006) runs until 2001 but has been extended. We thank Anthony Murphy for sharing this data. We get very similar results if we truncate the sample at 2001.

[16] Standard deviation computed from residuals of a regression of house-price-to-income on region dummies.

[17] There is also evidence that the credit conditions index matters for consumption (Aron et al, 2012).

[18] It is perhaps worth noting that, unlike the price variable, the credit conditions variable and the interest rate do not vary at the regional level. Additionally, and to anticipate, we find more of a role for credit conditions and the interest rate when we investigate changes in ownership rates after age 30 (see section IV).

[19] Full results of the analyses discussed in this paragraph are available on request from the authors.

[20] Full results of the analyses discussed in this paragraph are available on request from the authors.

[21] Of course, we could graph the level of ownership at age 40 (or 50) against ownership at age 30 (rather than having growth in ownership on the vertical axis.) This would eliminate the 2nd of the two biases described in the text. The slope of the graph would then be one plus the degree of catch up. Similarly, the regressions described below could be run with the level of ownership at age 40 (or 50) as the dependent variable, rather than the "growth regression" specification we currently use. This would allow the formulas suggested in Deaton (1985) to be applied directly. We work with ownership growth for two reasons. First, because of the convenient "catch up" interpretation of the slope (or slope coefficient), and second, because of the possibility that the two biases described in the text may partially cancel out. While corrections can be applied to the regressions below in either form, this seemed the best way to approach the data graphically.

[22] Descriptive statistics for these regressors are in online Appendix Table A.3.

[23] The catch-up coefficient estimates are robust to other ways of summarizing these variables, such as taking the average between age 31 and age 40, the growth rate etc.

[24] The mortgage payment to income variable is the ratio of average mortgage payment in the sample of 40 year-old owners, to average income in this sample. Averaging before taking the ratio improves robustness to the presence of "outlier" individual families with very high mortgage payment to income ratios (often due to low reported income).

[25] Full results with more regressors, and/or using two-stage least squares, are available from the authors.

[26] This, in our notation and in univariate form, is Deaton's equation (24) modified only for the fact that our x and y variables come from independent samples so we can continue to maintain the assumption that $\sigma_{uv} = 0$.

[27] We also experimented with using real mortgage payments at 40, and outstanding mortgage balances at 40, as outcome variables; however results (available on request) did not point to an advantage from early access to ownership, and coefficients were either not significant, or were not robust (even in sign) to the inclusion of regressors.