Housing, Stability and the Macroeconomy: International Perspectives

Federal Reserve Bank of Dallas
Journal of Money, Credit and Banking
International Monetary Fund

Rushing into American Dream? House Prices, Timing of Homeownership and Adjustment of Consumer Credit

Sumit Agarwal, National University of Singapore
Luojia Hu, Federal Reserve Bank of Chicago
Xing Huang, Michigan State University

This paper was presented at **Housing, Stability and the Macroeconomy: International Perspectives** conference, November 14-15 2013. The conference was sponsored by the Federal Reserve Bank of Dallas, the International Monetary Fund, and the *Journal of Money, Credit and Banking*. The conference was held at Federal Reserve Bank of Dallas (http://dallasfed.org).

Rushing into American Dream? House Prices, Timing of Homeownership, and Adjustment of Consumer Credit*

Sumit Agarwal[†], Luojia Hu[‡], Xing Huang[§]

Novermber 2013

Abstract

In this paper we use a large panel of individuals from Consumer Credit Panel dataset, and study the timing of homeownership as a function of credit constraints and expectations of future house price. Our panel data allows us to track individuals over time and we model the transition probability of their first home purchase. We find that in MSAs with highest quartile house price growth, the median individual become homeowners earlier by 5 years in their lifecycle compared to MSAs with lowest quartile house price growth. The result suggests that the effect of expectation dominates the effect of credit constraints and high price growth leads individuals to purchase home earlier. We further study other credit/loan behaviors around first-home purchases for young and old buyers. We find that younger buyers make more adjustments in their finances after the purchase– taking out more debt/credit, and yet they do not appear to experience larger increase in delinquency than older buyers.

Keywords: Consumption, Household Finance, Housing Wealth, Credit Constraints, Life Cycle

JEL Classification: D12, D14, D91, E21, E51, E62, G21, H31

*The authors would like to thank Anna Whittaker for excellent research assistance. We also benefited from conversations and comments from Zahi Ben-David, Souphala Chomsisengphet, Bo Honore, and Amit Seru. The views expressed in this research are those of the authors and do not necessarily represent the policies or positions of the Federal Reserve Board or the Federal Reserve Bank of Chicago.

[†]National university of Singapore - Departments of Economics, Finance and Real Estate

[‡]Federal Reserve Bank of Chicago

[§]Michigan State University - Department of Finance

1 Introduction

Homeownership is the ultimate American dream. Homeownership has several advantages to the household, the society, and the government as a whole. Homeownership increases neighborhood social capital investment because homeowners face high real estate transaction costs (reducing mobility) and have a financial incentive to increase their property value (see also DiPasquale and Glaeser (1999)). Additionally, households derive investment utility from the home. The society benefits because homeownership reduces crime, increases social interactions in the community, and builds neighborhoods. The government benefits because homeowners pay taxes and consumer at a higher level.

Modeling the demand for housing is complicated. Houses provide utility, serve as collateral for additional credit needs, and may bring investment benefits as well. This suggests that homeownership may vary over the lifecycle and over cohorts. Past literatures have focused on the impact of demographic variables (such as marriage), income and credit constraint. In this paper, we explore the influence of house price growth on the lifecycle of homeownership. We focus on first-home purchases, because first-home purchases account for 40% of sales over the past 30 years and more than 50% in 2009, according to the National Association of Realtors. First-home purchase may also matter for the long-term dynamics of housing market, because first-home purchase affects the demand for trade-up homes in the future housing market (Ortalo-Magne and Rady (2006)).

There may be two offsetting effects of house price growth on house demand. On the one hand, if price recently increased by a large amount, the individual may face a high level of house price. She has to sacrifice more consumption to pay the downpayment if she wants to buy it in the young period. When she is financially constrained, she will prefer to postpone the home purchase to her middle age period. Hence, house price growth decreases the probability of individuals to buy first-home in early ages. On the other hand, individuals may extrapolate future house price growth from past house price growth. They are more likely to buy a house early in their life cycle if they expect the house prices to rise faster. In this case, house price growth increases the probability of individuals to buy first-home earlier.

In this paper, we empirically test the impact of house price growth and study which effects dominate in the data. We exploit a large panel of individuals from the FRBNY Consumer Credit Panel (CCP) dataset. We follow each individual from 1999 to 2012 and study their homeownership timing decision as a function of house price changes across MSAs controlling for other demand and supply variables. This data set is truly unique to study this question because of the following three reasons: (i) The data is a long panel of quarterly

information about limited demographics and risk measures that are time varying with little measurement errors (unlike the survey datasets that can be potentially biased, e.g. Survey of Consumer Finance); (ii) The data has detailed geographic information up to the zip code level; (iii) Unlike surveys datasets, this is a panel of the entire population that are active in the credit markets and so our estimates are reasonably precise. The strength of the data allows us to exploit the time-varying cross-regional variation in house price growth and its impact on the timing of homeownership.

Despite the long panel nature of the dataset, 13 years is not sufficient to study lifecycle decisions. For instance, it is possible that even after 13 years many individuals may not have made a decision to buy a house, so as an econometrician we face a right censoring problem. To deal with the censoring issue, we estimate the transition into homeownership using a discrete time survival model. This also allows us to study the entire age distribution of the potential home buyers (as some will never buy a home).

Our main variable of interest is the MSA level house price change which is a proxy for expected future house price growth. According to a survey of US home-buyers in 2002 conducted by Case, Quigley, and Shiller (2003), home-buyers expectations are substantially affected by recent experience. In the current analysis, we assume households in the same region hold same expectations. Their expectations are all based on the past three-year house price changes. In the future analysis, we could relax this assumption to allow households form expectations on their experiences in their lifetime[1], which means that expectations could vary across age groups within a region.

There could be confounding factors that could explain the timing of homeownership decision. To deal with this, we have MSA-fixed effects to control for level of house prices across MSAs and credit supply that is not time varying. This also controls for some demographic and other market conditions. We also have time-fixed effects that control for the variation in demand and supply of credit and house prices at the aggregate level. We have MSA-specific time-varying control variables like unemployment rate, growth rate of number of businesses, and wages. We also control for individual-specific time-varying credit shocks by their credit risk scores (this measures their ability and willingness to pay credit).

To fix ideas, the thought experiment that we have in mind is that we take two random people and assign them to the highest and lowest house price growth MSA quartiles and assume that they do not move between MSAs. Then we study if the house price growth in the quartiles affects their timing of home purchase. We show that individuals accelerate

[1] Malmendier and Nagel (2011), Malmendier and Nagel (2013) find that households form stock returns and inflation forecasts based on their experiences on stock returns and inflation in their lifetime. Similarly, when households forecast future house prices, they may take into account their experiences in their lifetime. Hence, the forecasts of future house prices may also be age dependent.

their probability of buying a house in MSAs in the highest quartile of house price growth relative to MSAs with the lowest quartile of house price growth. Specifically, we find that the median age of first home buyer in the population goes down (buy at an earlier age) by 5 years when we compare individuals that live in the highest house price growth quartile to the lowest house price growth quartile. However, conditional on individuals who have eventually bought a home by age 60, the same difference is only 1 year. Our interpretation of this result is that individuals have a higher expectation of future house price appreciation and this leads them to purchase home earlier.

Next, we study whether the shift of lifecycle of homeownership would affect individuals' other credit behaviors: since younger buyers may be more financially constrained when they purchase the houses, would their other credit accounts experience a larger increase of borrowing relative to the accounts of older buyers? If so, would the increase of default risk for younger buyers be larger as well? Home purchases would potentially have significant impact on these behaviors, because it is highly levered and probably the largest single investment of the household. Our findings suggest that younger buyers make more adjustments in their finances after the purchase– taking out more debt/credit, and yet they do not appear to experience larger increase in delinquency than older buyers.

Our paper is most closely related to Landvoigt (2011), in the sense that he also studies the role of credit constraints and house prices expectation on the size of home purchased and the decision to purchase. However, there are some key differences between the two studies. First, we focus on the timing of homeownership as opposed to decision to buy a house. Second, we proxy for expectations on future house prices using past house price changes and investigate the effects of expectations. Differently, Landvoigt (2011) aims to infer house price expectations from observed household choices. Furthermore, we use an administrative panel dataset with little measurement error of individuals with detailed geographic information that allows us to exploit the geographical variation in house price growth.

Our paper is related to the large literature that studies the life-cycle housing demand. Ortalo-Magne and Rady (2006) study a life-cycle model of the housing market with a property ladder and a credit constraint. Different from their emphasis on the link between income shock and house prices, our paper focuses on the influence of house price expectations on the behavior of first-time home buyers. Attanasio, Bottazzi, Low, Nesheim, and Wakefield (2012) also constructs a life-cycle model and incorporates some realistic features, but they are still absent of one feature considered in our paper that households could extract time-varying utility from home ownership. Sinai and Souleles (2005) model the demand for owning as the trade-off between the rent risk and the asset price risk. They relate the demand for homeownership to local rent volatility and households expected horizon. Han (2010) identi-

fies two effects of price risk on housing demand: a financial risk effect and a hedging effect against future housing costs. The author studies the timing and size of house purchases by existing homeowners. Our paper differs by studying the timing of the marginal first-time home buyers.

Our paper also contributes to the growing literature that finds evidence linking the creation of the real-estate bubble in the early 2000s to demand for mortgages – Keys, Mukherjee, Seru, and Vig (2010), Ben-David (2011, 2012), Berndt, Hollifield, and Sandas (2010), Agarwal, Ben-David, and Yao (2012), and Jiang, Nelson, and Vytlacil (Forthcoming). Our second part also relates to recent literature discussing household's credit behavior – Agarwal, Driscoll, Gabaix, and Laibson (2009), Debbaut, Ghent, and Kudlyak (2013), Sullivan (2008), Karlan and Zinman (2010), Morse (2011).

The paper proceeds as follows. Section two goes through the datasets we use and summary statistics. Section three forms the hypothesis and tests the impact of house price growth on the timing of first-home ownership. Section four further studies the adjustment of other credit behavior around first-home purchases, and section five concludes.

2 Data and Summary Statistics

2.1 FRBNY (Equifax) Consumer Credit Panel

To track individual mortgages and credit history, we use a 1% sample of the primary-ID level Equifax dataset. This is a panel dataset collected each quarter from 1999-2012. Individuals in the panel are selected randomly from the US population based on the last two digits of their social security number. Then the FRBNY collects credit bureau data for these individuals, including mortgage and non-mortgage debt, collection agency records, and personal background information.

Individuals can only be selected into the sample if they have a credit record on file that includes their social security number. This means that as soon as a young adult with a randomly selected SSN opens his or her first line of credit (often around age 18), he or she will be added to the Equifax dataset. Deceased individuals are also dropped from the dataset. This sampling methodology should ensure that the random sample reflects the current demography of the US population with a credit history and social security number.

In practice, we only include individuals age 18-60 in the data since elderly individuals may have already paid off their mortgage long enough in the past[2] that it no longer appears in

[2]According to the FRBNY staff report listed above: Closed accounts remain on credit reports for up to 7 to 10 years after their closing. Therefore, our panel includes those with no recent credit activity, such as in the past 24 months, but with credit activity in the past 10 years (footnote 4, pg 2). However, detailed information on specific accounts (such as a mortgage) must be updated by an individuals creditors in the

their credit file. Credit files for the very old and the very young are often small or incomplete, so we exclude them from our analysis.

The key variables we select from the Equifax dataset are a persons age, address, credit score, and mortgage history. Table 1 reports summary statistics. If we pool all the individual-year observation, the average age of the sample is about 40. We only keep the individuals between 18 and 60. The credit score ranges from 291 to 842 with a mean 666.9,

We are especially interested in each persons oldest mortgage account, which would indicate their first home purchase. Since we want to analyze home purchase decisions at a yearly frequency, we select each persons age, credit score and address at the beginning of each year (usually Q1, unless an individual enters the dataset mid-year). We then look to see if they purchased their first home at any point during that calendar year. As Table 1 shows, the average age to purchase the first-home is 35.4 in our sample.

We want to point out a few shortcomings of the data selection and exclusion process. About 5% of individuals in the Equifax sample had missing birth years, so we could not identify their age. Because we want to examine differences in home purchase decisions across different age groups, we dropped anyone with a missing birth year. To the extent that missing age data occurs randomly, this should not affect our analysis.

We also want to keep track of each persons location each year, but individuals frequently change their address over time in the Equifax data. Equifax collects the primary billing address listed by each persons creditors, which should normally correspond to their physical address, but this seems to be a noisy measure. In some specifications we exclude individuals who move from one city to another during a given year.

We also found it difficult to accurately measure the age of a persons oldest mortgage account. This variable should stay constant for each individual over time as they remain in the survey; however, this was not always the case. Very often, Equifax would report a given age of oldest mortgage for an individual at one point in time, but they would report a different value at a later point in time during the panel. We decided that we would always take the age of oldest mortgage from the earliest survey response where a person had reported having ever taken out a mortgage. We then replaced any subsequent oldest mortgage values to match the oldest non-missing survey response.

2.2 CoreLogic House Price Index

We use CoreLogic home price index (HPI) data to compare housing price growth in different regions as a measure of expected growth in future housing prices. CoreLogic home

past 3 months in order to be included in the Equifax dataset for a given quarter: While a record will be included in our panel for all individuals with some credit activity on their credit reports over the past 7 years or sorecords will only include information on recently updated accounts (footnote 14, pg 10).

price indices are calculated using weighted repeat sales methodology on a monthly frequency, with January 2000 as the base month. For our analysis we select data only for single family combined homes (including distressed sales). For these homes, we use HPI data calculated at the CBSA-level to capture variation in price growth between different metropolitan areas across the country.

To compare changes in housing prices, we use CoreLogic data which creates a housing price index measuring variation in the CBSA-level (as well as national-level) housing prices from the base year (2000). First, we calculate the average HPI at year t in city c as the 12-month average of the monthly HPI from CoreLogic. We then compute the annual growth rate of HPI from year $t-1$ to year t as[3]:

$$\gamma_{c,t} = \frac{p_{c,t} - p_{c,t-1}}{p_{c,t-1}}$$

We smooth the house price growth by taking the average of HPI growth over the most recent three years:

$$\delta_{c,t} = \frac{1}{3}\sum_{k=0}^{2} \gamma_{c,t-k}$$

Over the sample period and across all CBSAs, the average annual HPI growth rate is about 3% a year (Table 1). There is variation both across cities and over time. We compute the range of the HPI growth over the sample period (taking the difference between the highest- and the lowest- growth rate) for each CBSA. According to Table 1, within a given CBSA, HPI growth varies over time the average range across all CBSAs is 21 percentage points. More importantly, this range also varies quite a bit across CBSAs. For example, while the annual HPI growth rate changed by no more than 13 percentage points over the period in one-fifth of the cities, it swung by 27 percentage points or more in another one-fifth of the cities.

2.3 BLS: Employment and Wage Data

We collect regional employment and wage data from the Quarterly Census of Employment and Wages (QCEW), compiled by the BLS. This is a near-census of business establishments across the country. At the county level, we select annual average employment levels, weekly wages, and number of establishments and then calculate growth rates for the each of these variables.

In addition to the QCEW data, we use county-level unemployment rates from the Local

[3]For robustness, we also experimented with using the growth rate from Quarter 4 of year $t-1$ to Quarter 4 of year t, and the results remain.

Area Unemployment Statistics published on the BLS website for 1999-2012.

Since these variables are given at the county-level, we map each county to its corresponding CBSA in order to link it with the housing price data. If a CBSA includes more than one county, we take an average of these four macroeconomic indicators across all counties within a given CBSA.

3 House Prices and the Timing of First Homeownership

3.1 Hypothesis

We will explore how house price growth affects the housing demand over the life-cycle using micro level data. To form the hypothesis, we provide a conceptual model in Appendix A.1. Assume households preferences for owning a house varies over the life cycle. A household could obtain huge extra consumption utilities from homeownership in her middle age relative to in her young age. House price growth could affect the housing demand through two channels.

The first channel is through *liquidity constraints*. If price recently increased by a large amount, the individual may face a high level of house price. As price increases, all else equal, she has to sacrifice more consumption to pay the downpayment if she wants to buy it in the young period. When she is financially constrained, she will prefer to postpone the home purchase to her middle age period. Hence, house price growth decreases the probability of individuals to buy first-home in early ages.

The second channel is through *house price expectations*. Our model also predicts that housing demand will vary with the house price expectations. Specifically, households are more likely to buy a house early in their life cycle if they expect the house prices to rise faster. The intuition is as follows. When the expectation of future house prices is low, the household would prefer to postpone their purchase of the first home. Otherwise, she has to sacrifice her consumptions in her young age for the downpayment while without experiencing as many utilities from homeownership as in her middle age. In contrast, when the expectation of future house prices is high, the household would move her first-home purchase earlier. There are two incentives to do so. The first incentive is investing. She could obtain the potential capital gain from the house purchase. The second incentive is hedging. She could hedge against the possibility that she may not be able to afford the downpayment in her middle age and hence cannot enjoy the extra utilities from homeownership, if the price reaches sky high. Through this channel, house price growth increases the probability of individuals to buy first-home earlier.

3.2 Homeownership Rate

The Equifax data does not explicitly measure whether an individual owns a home at the survey date or when an individual bought his or her first home. It does, however, record the age of their oldest mortgage account. Based on this information, we derive a measure for whether or not an individual ever owned a home in a given year, and among those who did, the age at which they bought their first home.

For the purpose of this paper, we consider an individual to be a home owner in a given year t if he or she purchased a home in year t *or earlier*. Accordingly, we define the home ownership rate in year t as the fraction of individuals in year t who have ever purchased a home by year t. Specifically, let Bt denote the total number of individuals who bought a home in year t or earlier, and N_t denote the total number of individuals at year t. Then the home ownership rate in each year, H_t, can be calculated as the ratio:

$$H_t = \frac{B_t}{N_t}$$

We calculate the home ownership rates by age in a similar fashion. Specifically, let Bat denote the total number of individuals in age group a at year t who purchased a home in year t or earlier, and N_{at} denote the total number of individuals in age group a at year t. Then the homeownership rates for age group a in each year, H_{at} , can be calculated as the ratio:

$$H_{at} = \frac{B_{at}}{N_{at}}$$

Figure 1 plots the aggregate home ownership rate H_t over the time period 1999-2012. Over the sample period between 1999 and 2012, the fraction of home owners varies between 44% and 47%, with an average around 46%. This is considerably lower than the homeownership rate measured using other data (such as the Census, CPS or AHS). For example, Fisher and Gervais (2011) use the census data to examine homeownership trends and found that over 1996-2007, the homeownership rate for all individuals age 25+ rose from roughly 65% to 68%.

[Insert Figure 1]

There are several possible explanations for why our findings differ from the literature [4], including differences in the data and the definition of home ownership. For example, Fisher and Gervais (2011) use census data for individuals ages 25+ and they measure the homeownership rate as the number of owner-headed households divided by the total number

[4] Our finding, however, is similar to Mian and Sufi (2011). Using the same Equifax data but with somewhat different definition, they also found the home ownership rate to be around 40%.

of occupied households. In contrast, our dataset is based on mortgage data from a random sample of all individuals 18 to 60 years old with a credit report. By including people younger than 25 (who do not usually own homes) and excluding the elderly (who often do) it is not surprising that we find a lower fraction of homeowners in our sample.

Moreover, our definition of home ownership is somewhat non-standard. We consider any individual who has ever opened a mortgage account as a homeowner even though he/she might not still own a home at the survey date. We also exclude actual owners if they have never had a mortgage or paid off the mortgage account more than 7-10 years ago[5] (also see Mian and Sufi (2011)).

Figure 2 plots the home ownership rates for three age groups (age 18-24, 25-34, 35-44, and 45-60) against the 3-year average house price growth for the sample period. For the young group from age 25 to 34, the fraction of individuals ever owning a home positively comoves with the recent house price growth. In contrast, the relationship between homeownership for the other two older age groups (35-44 and 45-60) and house price growth seems weaker - homeownership increases as the house price growth increases in early 2000s, but stays stable from 2004 to 2008 and only starts to fall after 2008. The different relationships may suggest that house price growth may matter more to the decisions of the young individuals than to the decisions of the older ones. In the next section, we examine whether and how (expected) house prices affect peoples decisions about when to enter the housing market.

[Insert Figure 2]

3.3 Hazard Rate of First Home Purchase by Age and Housing Price Growth

For the purpose of our paper, the key concept to examine is the hazard rate of first home purchase. In other words, what is the probability that an individual will buy a home in a given year, conditional on the fact that they never bought a home in a previous year? We do this by truncating our dataremoving individuals from the dataset the year after they purchase their first home.

Using cross-sections from this data, we can compute the hazard rate (conditional probability) of first home purchase at a given age by year and geographic area. Let $\tilde{H}_{ct}(a)$ denote the total number of individuals of age a living in area c at year t who purchased their first home in year t, and $\tilde{R}_{ct}(a)$ denote the total number of individuals of age a living in area c

[5] Once a mortgage is paid in full, the account is listed as a closed account on the individuals credit report. Closed accounts do not usually stay on a credit report for more than 7-10 years.

at year t who have never bought a home before year t, then the hazard rate is given by:

$$\tilde{h}_{ct}(a) = \frac{\tilde{H}_{ct}(a)}{\tilde{R}_{ct}(a)}$$

We first compute the hazard rate of home purchase by age for each city-year cell as described above. Then we group the city-year cells into five groups based on HPI growth, and compute the average hazard rate across city-year cells within each group, weighted by the cell size (i.e. the number of people in the city-year cell). Figure 3 presents the average hazard for each of the five groups, ranging from the lowest- to the highest-quartile of housing price growth.

Figure 3 shows that the hazard rate of home purchase over the lifecycle is hump-shaped it increases sharply after the mid-20s, peaks in the early 30s and declines afterwards.

The figure also shows that at a given age, the hazard rate of home purchase in cities during periods of faster house price growth is generally higher than that in cities during periods of low price growth. Moreover, the gap widens at ages between the mid-20s and the early 30s, and stays roughly constant afterwards. This suggests that people tend to purchase their first home at younger ages when they live in cities during periods of fast price appreciation.

To look more directly at the house price effect on the distribution of age-at-purchase, we can use the estimated hazard to compute and compare the counterfactual CDFs under the assumption that an individual always lives in the low- versus the high-house price growth area.

First note that with the estimated the hazard rate (conditional probability) of home purchase for each age, we can also compute the corresponding unconditional probability of purchasing a home at or *before* a given age (i.e. the cumulative distribution function). Specifically, for each age (j) from 18-60, we calculate the probability that an individual would purchase a home at or before that age as :

$$\tilde{F}_j = 1 - \Pi_{18}^{j}(1 - \tilde{h}_j)$$

For example, the probability of buying a home by age 20 is given by

$$\tilde{F}_{20} = 1 - (1 - \tilde{h}_{18})(1 - \tilde{h}_{19})(1 - \tilde{h}_{20})$$

Based on this, we first compute the CDF by city-year cell and then take the (weighted) average across all cities and years within each of the five HPI growth groups. Figure 4 presents the resulting (counterfactual) CDFs.

Figure 4 shows that the distribution of the age at first home purchase for the higher HPI growth city-year groups generally lies uniformly above and to the left of the distribution for the lower HPI growth city-year groups. In other words, the former distribution stochastically dominates the latter. Figure 3 and 4 suggest that people living in cities during periods of high house price growth are generally more likely to purchase their first home at a younger age than those living in places with low house price growth. The result is more consistent with the house price expectation channel instead of the liquidity constraints channel.

Note that so far our comparisons are made between city-year groups with high versus low housing price growth and used only the aggregate cross-section data. As such, we cannot distinguish whether the differences in home purchases by age come from variations in price growth across cities in a given year or from variations over time within a city. To sort these out, and to also take into account other factors that might affect an individuals decision to purchase a home, we turn to multivariate analyses using individual-level panel data in the next section.

3.4 Estimation of Hazard Rate – Multivariate Analysis

We use a simple discrete-time hazard specification to model the probability of first home purchase, using the Equifax individual-level data. Specifically, we use a Logit model for the binary outcome of home purchase (conditional on never previously owning a home). The latent variable is

$$y^*_{iact} = \beta_1 \cdot \Delta HPI_{ct} + \beta_2 \cdot Age_a + Year_t + CBSA_c + X'_{it}\eta_1 + M'_{ct}\eta_2 + \varepsilon_{iact}$$

where i indexes individual, a indexes age, c indexes CBSA and t indexes year. The model includes the average HPI growth rate over the past recent years ΔHPI_{ct}, year- and CBSA-fixed effects, single-year age dummies, individual-level time-varying variable X'_{it} (risk-score), and variables proxying the local economic conditions M'_{ct} (e.g. growth rates of the number of businesses, employment and wage, and the unemployment rate). The hazard rate for an individuals first home purchase is then

$$\begin{aligned} h_{ict}(a) &= Pr(Buy_{iact} = 1 | Buy_{iac\tau} = 0,\ \tau < t) \\ &= Pr(y^*_{iact} > 0) \\ &= \Lambda(\beta_1 \cdot \Delta HPI_{ct} + \beta_2 \cdot Age_a + Year_t + CBSA_c + X'_{it}\eta_1 + M'_{ct}\eta_2) \end{aligned}$$

where $\Lambda(\cdot)$ is the logistic CDF $\Lambda(u) = \frac{e^u}{1+e^u}$.

Table 2 below gives results from the full sample of CBSAs from 1999-2012 using the Equifax primary-level data (1% sample). We included all individuals 60 and younger.

[Insert Table 2]

Our main result can be found in Column 1. The coefficient on HPI growth is positive and significant. Assuming individuals form their expectations of future house prices based on the recent housing price growth, the result suggests that all else equal, at a given age, individuals who live cities with higher expected future house price appreciations are more likely to buy the first home than their counterparts in areas with low growth in expected housing prices.

There is distinct lifecycle pattern in the hazard of first home purchase. The coefficients on the age dummies (not reported here) exhibit a skewed hump-shaped age profile that is similar to what we saw before the hazard rises sharply from the late 20s, peaks around 30 and then declines gradually afterwards. Given this pattern, house price expectations might have a differential impact on the likelihood of first home purchase at different ages.

There are many other factors that could affect whether an individual is willing and/or able to buy a home. For example, as most home purchases are financed by a mortgage, and having a good credit score is crucial for obtaining a loan, we would expect credit scores to have an important effect on home purchases. And this is exactly what we found. The risk score variable enters the model positively and with strong statistical significance.

Our hypothesis is based on the assumption that we can interpret the recent house price growth as a proxy to future house price growth. One potential problem is that housing price growth might also be correlated with other economic conditions that affect housing demand, regardless of the expectation of future prices. For example, cities that experienced rapid house price appreciations might also have had fast growing local economies with more jobs and increasing wages. Higher incomes, in turn, could make a house more affordable and thus lead to higher and earlier entry into the housing market. To address this concern, we added some CBSA-level controls to capture the time-varying local economic conditions. Specifically, we include the growth rates of employment, wages, and business establishments as well as the unemployment rate. Columns 2 and 3 show that while local employment, wage, and business growth seem to have little additional impact on the likelihood of an individuals first home purchase, the local unemployment rate has a negative impact on the likelihood of first-home purchase. While the effect of HPI growth on the hazard rate is reduced somewhat with the addition of each variable, it remains statistically significant even in the full model (Column 3).

We also experiment with a specification that relaxes the linearity functional form assumption on the housing price variable. Specifically, we replace the continuous variable, HPI growth, by a set of dummies that represents the quartiles of its distribution. The results in Column 4 show that generally the hazard of home purchase indeed increases monotonically

with HPI growth.

Since the 3-year HPI growth might also reflect the growth from $t-1$ and t, the period may have overlap with the time of home purchase. One concern is that increases of home purchase may cause house prices to go up during year t, leading to reverse causality. To resolve this concern, we also use another measure, 2-year HPI growth from year t-3 to t-1, which does not include the current year. The result (Column (5)) remains the same. It implies that our result is not driven by the effect of increasing house demand on the current house price.

We are also concerned about possible endogeneity of housing price growth at the local level, so we create an instrument for our 3-year change in HPI growth variable. The instrument is equal to a national measure of house price growth each year (also taken from CoreLogic) multiplied by the elasticity of housing supply in each CBSA[6]. We then run an ivprobit regression. The results for the first and second stage are listed in Appendix Table A.1, which continue to show a positive effect of housing price growth on the likelihood of home purchase.

Another concern is the left-censoring problem. Because our data on individual credit scores, addresses, etc. only goes back to 1999, we have to limit our analysis to the years 1999-2012. This may cause a left-censoring problem, since we cannot observe past data for older individuals who have never bought a home when they enter our data in 1999. We perform a robustness check in Appendix, only keeping young individuals (22-25 years old when they enter our data in 1999). Since this cut our sample size dramatically, we used a 5% Equifax primary random sample, rather than the 1% primary sample we used for the rest of our analysis. Our results remained the same, as shown in the Table A.2.

3.5 Evaluate the Magnitude of the HPI Effect

While the estimated coefficients from the model can show that HPI growth has a positive effect on the hazard of an individuals home purchase, the magnitude of this impact is not immediately clear since the model is highly nonlinear. Moreover, since the hazard rate is only a conditional probability, it might not be the final object of interest if, for example, one wants to answer questions such as "if house prices increase by 10 percentage points, how big an increase would there be in the share of individuals who have bought a home by age 30?".

In this section, we conduct two counterfactual experiments. We consider two scenarios: (Case 1) we assume that individuals in our sample have always lived in the cities with the

[6]Saiz (2010) calculates these elasticities for 95 of the largest CBSAs, so we include elasticity data for these CBSAs (which cover a little less than half of the observations in our survey data). The elasticities from Saiz (2010) are only measured at one point in time, so we assume that they do not change over the sample period.

lowest HPI growth (the bottom quartile) vs. (Case 2) we assume that individuals in our sample have always lived in cities with highest HPI growth (the top quartile).

[Insert Figure 5]

Under each scenario, we use the estimates from the Logit model in Column 4 of Table 2 to predict for each individual the hazard of home purchase at each age between 18 and 60. For each prediction, we only vary the HPI growth and age variables at their hypothetical values and keep all other variables at their actual values in the data. We then take the average of the predicted hazard across all individuals. Figure 5 Panel (a) presents the average hazard for the top quartile and bottom quartile separately. Panel (b) presents the difference. The figures show a heterogeneous effect of house price growth. The effect varies by age group - when house price growth increases, the hazard rate of buying first-home increases by a larger amount among young individuals.

Similarly, under each scenario, we also use the predicted hazards to estimate the CDF for each individual at each age and then take the average across all individuals. Figure 5 Panel (c) presents the average CDF from the experiment.

According to the counterfactual distributions in Panel (c), the difference in the median age of first-home buyers (which is defined as the age by which half of the population has bought a home) is about 5 years: first-home buyers are 5 years younger under scenario 2 than under scenario 1 (39 vs. 44 years).

Note that as we only have estimates of the hazard and survivor functions for individuals aged 18 to 60, we can not reliably estimate the expected value (mean) of age-to-purchase for the entire population. The reason is there are many people in the population who will never buy a home. However, with the available estimates, we can still calculate some summary measures for the *conditional* distribution of age-at-purchase among those who will have eventually bought a home by age 60.

Specifically, we can estimate for each individual i the conditional mean age as:

$$\hat{E}(a_i|a_i \leq 60) = \frac{\sum_{s=18}^{60} s\hat{h}_i(s)\hat{G}_i(s-1)}{1-\hat{G}_i(60)}$$

where $\hat{G}_i$ is the survivor function, defined as $\hat{G}_i(t) = \Pi_{s=18}^{t}(1-\hat{h}_i(s))$.

Based on this calculation, we find that among those who have bought a home by age 60, the average age at first home purchase is about half year younger under scenario 2 than under scenario 1 (33.7 vs. 34.2 years).

Finally, we can also estimate for each individual i the conditional distribution function

as:

$$\hat{F}(a_i|a_i \leq 60) = \frac{\hat{F}(a_i)}{1 - \hat{G}_i(60)}$$

where

$$\hat{F}_i(t) = 1 - \Pi_{s=18}^{t}(1 - \hat{h}_i(s))$$

or equivalently

$$\hat{F}_i(t) = \sum_{s=18}^{60} \hat{h}_i(s)\hat{G}_i(s-1)$$

Based on the estimated conditional CDF, we find that the median age at first home purchase among those who have eventually bought a home by age 60 is about 1 year younger under scenario 2 than under scenario 1 (31 vs. 32 years). Note the difference in the conditional median is smaller than the difference in the unconditional median we estimated earlier.

4 Do Young and Middle Age People Adjust Differently Around First-home Purchases?

Since home purchase accounts for a large investment in most individuals' portfolio, the shifting of the life cycle of homeownership may have further implications on individuals other investment behavior. Our dataset allows us to observe individuals' consumer credit behavior, such as bank credit card accounts, auto loans, student loans, etc., as well as home equity-based borrowing behavior, such as home-equity installment loan and home-equity line of credit. In this section, we turn to explore how these behaviors would change around/after first-home purchases, and especially whether the adjustment would be different between younger and older individuals. Large adjustment of other credit accounts may further affect the delinquency rates of these accounts. We'll also discuss the change in delinquency rates of various accounts after first-home purchases, and compare between younger and older individuals.

4.1 Event Study

We use event study to estimate the change of individuals' credit behaviors around their first-home purchases. Following Jacobson, LaLonde, and Sullivan (1993), we estimate the following specification:

$$
\begin{aligned}
Y_{iact} &= \alpha_i + \sum_{k=-3}^{6} \beta_k \cdot \mathbf{1}_{t-\tau^*=k} + \sum_{k=-3}^{6} \eta_k \cdot \mathbf{1}_{t-\tau^*=k} \cdot \mathbf{1}_{a(\tau^*)\geq 35} \\
&+ Age_a + Year_t + CBSA_c + \varepsilon_{iact}
\end{aligned}
$$

where Y_{iact} denotes the outcome variables, including credit score, credit/loan balances in various accounts, and delinquency rates. i indexes individual, a indexes age, c indexes CBSA and t indexes year. Dummy variables, $\mathbf{1}_{t-\tau^*=k}$, are event time indicators, which equals to 1 if year t is k years away from the first-home purchase year τ^*[7]. To compare the changes between younger and older individuals, we interact the event years with a dummy variable, $\mathbf{1}_{a(\tau^*)\geq 35}$, which is 1 if individuals are older than (or equal to) 35 years old[8]. We also include (1) the individual effect, α_i, allowing for arbitrary permanent heterogeneity among individuals in unobserved characteristics; (2) the age effect, Age_a, capturing the average age pattern of credit behavior; (3) the year effect, $Year_t$, controlling for the influence of macroeconomic factors; (4) the city effect, $CBSA_c$, identifying the heterogeneity among cities.

Coefficients β_k captures the change of the outcome variables for younger individuals relative to their own past, controlling for age, year and city effects. Accordingly, $\beta_k + \eta_k$, provides the change of the outcome variables for older individuals relative to their own past, also controlling for age, year and city effects.

4.2 Adjustment of Consumer Credits

We first look at the adjustment of the credit accounts not related to home equity, for example, bank credit cards, auto loans, and etc. The coefficients β_k and $\beta_k + \eta_k$ are plotted in Figure 6. From Panel (a), (b) and (c), we find that both groups take out more debt (in credit cards, auto loan or retail cards) around their first-home purchases. The balances sharply increase during the year of the home purchase, and do not revert afterwards. As for student loans (Panel (d)), we find both groups slow down their repayment relative to their cohorts without home purchases. Besides, younger and older individuals have significantly different behaviors in their credit card accounts, auto loans and student loans (Panel (a), (b), and (d)), but the difference in their retail and other accounts is insignificant (Panel (c)). Younger individuals live upon more credits in their credit cards and auto loans, and repay their student loan more slowly, relative to older individuals.

[7]For notation convenience, we summarize all the years more than 6 years away from the first-home purchase year into $k = 6$. So, $k = 6$ stands for $t - \tau^* \geq 6$.

[8]We set the cutoff as age 35, because we find that the hazard rate of purchasing the first-home increases more for individuals younger than 35 relative to the increases for individuals older than 35. We also experiment with other cutoffs around 35, and the results remain the same.

We also study the change in their mortgage accounts and home equity-based borrowing (Figure 7). For this analysis, the sample only includes individuals who ever bought a house in our sample. As individuals pay down the mortgages and home-equity installment loans, the balances decrease through years (Panel (a) and (b)). As Panel (a) shows, even though older individuals may own a bigger house, they start with a smaller loan, which may indicate a larger downpayment. Besides, older individuals pay down mortgages faster than younger ones. As for home-equity installment loan, it seems older individuals owe a larger amount compared to younger ones. Given installment loans usually are used for house improvement, a bigger house owned by older individuals may lead to a larger loan for the improvement. Differently, younger individuals borrow significantly more money from home-equity line of credit, which is consistent with their other credit behaviors, such as credit cards and auto loans.

In all, the figures show that there seems to be a shock to individuals' credit behavior around first-home purchases. Individuals with home purchases borrow more than their past selves and their peers without home purchases. Especially, the shock is larger for younger individuals than for older ones - younger buyers borrow a larger amount and pay down debt slower than older buyers.

4.3 Change in Credit Score and Delinquency Rate

Then we turn to examine the default risk of these individuals and how it may change around their first-home purchases. Figure 8 presents the change in credit score as a measure of individuals' default risk. The credit scores for home buyers gradually increase before their home purchases and revert after home purchases. In the long run, the scores are still higher than the level three years before home purchases. This could come from individuals with increasing credit scores self-select to purchase houses. It is also possible that some individuals build up their credit scores right before purchasing houses. In any case, we observe a long run increase of credit scores around home purchases, and the increasing amount is larger for younger buyers than for older buyers.

We also directly look at the default behavior of these individuals. The variable we use is the severe delinquency (more than 90 days past due) of various accounts. The observations only include individuals who have a positive balance on the account. Figure 9 shows the delinquency rate for the credit accounts not related to home equity. Generally speaking, the delinquency rate of other consumer credit accounts increases after the home purchases, and it keeps rising as years go by. Moreover, for credit cart accounts, auto loans and student loans, the delinquency rates of older buyers seem to increase more compare to those of younger buyers (Panel (a), (b), and (d)). The delinquency rate of retail or other accounts is not

significant different between these two groups (Panel (c)).

The delinquency rates of mortgage accounts and home equity-based borrowing are displayed in Figure 10. Consistent with other credit behavior, the delinquency rates increase through years after the home purchases. But there are not significant differences between younger buyers and older buyers.

5 Conclusion

In the paper we investigate how house price growth affects individuals home purchase over the lifecycle. In particular, we focus on the role of expectations about future house prices on the timing of entry into the housing market. We use a unique panel dataset from 1999-2012 that allows us to track the same individual over these 13 years and observe their decision to buy a house. We show that when the house prices are rising, individuals have a higher house price appreciation expectation and tend to buy houses earlier in their lifecycle. This is possible due to the availability of credit in these markets. To separate from the confounding effects of time-varying credit constraints, we also exploit the variations in expectations across regions. We show that households accelerate their probability of buying a house in MSAs with highest quartile house price growth relative to MSAs with the lowest quartile house price growth. Specifically we find that the median age of first home buyer in the population goes down (buy at an earlier age) by 5 years when we compare individuals that live in the highest house price growth quartile to the lowest house price growth quartile. We also discuss how the shift of lifecycle of homeownership affects individuals other credit behaviors. We find that younger buyers make more adjustments in their finances after the purchase – taking out more debt/credit, and yet they do not appear to experience larger increase in delinquency than older buyers.

Our key contributions to the existing literature can be summarized as follows. First, we are the first to use the FRBNY CCP dataset to study this question. This dataset is a large panel, with little measurement error, following individuals over a relatively long time horizon with detailed geographical information. We also use a survival type of analysis that explicitly deals with the right censoring problem. This also allows us to study the entire age distribution of the potential home buyers (as some will never buy a home). Second, the prior literature mostly focuses on the level of home ownership and not the timing of home ownership. We show that the timing is also impacted by the relaxation of credit constraints and the regional variation in (expected) future house prices growth - individuals pre pone the house purchase decisions when they have high expectation on future house prices. This may have significant policy implication for the CFPB on how to help the young in the house buying decisions.Third, we study the impact of home purchases on other credit

behavior. The findings suggest that though younger buyers increase their debt/credit by a larger amount after first-home purchases, and yet they do not appear to experience larger increase in delinquency than older buyers. This is consistent with the findings by Agarwal, Driscoll, Gabaix, and Laibson (2009). Our findings may have implications on how the market should charge for the default risk of young buyers vs older buyers.

References

Agarwal, Sumit, Itzhak Ben-David, and Vincent Yao, 2012, Collateral valuation and institutional pressures: Evidence from the residential real-estate market, *Working Paper*.

Agarwal, Sumit, John C. Driscoll, Xavier Gabaix, and David Laibson, 2009, The age of reason: Financial decisions over the life-cycle with implications for regulation, *Brookings Papers on Economic Activity* Fall.

Attanasio, Orazio P., Renata Bottazzi, Hamish W. Low, Lars Nesheim, and Matthew Wakefield, 2012, Modelling the demand for housing over the life cycle, *Review of Economic Dynamics* 15, 1–18.

Berndt, Antje, Burton Hollifield, and Patrik Sandas, 2010, The role of mortgage brokers in the subprime crisis, *Working Paper*.

Case, Karl E., John M. Quigley, and Robert J. Shiller, 2003, Home-buyers, housing and the macroeconomy, in Anthony Richards, and Tim Robinson, ed.: *Asset Prices and Monetary Policy* . pp. 149–188 (Canberra: Reserve Bank of Australia).

Cocco, Joao F., 2005, Portfolio choice in the presence of housing, *Review of Financial Studies* 18, 535–567.

Debbaut, Peter, Andra Ghent, and Marianna Kudlyak, 2013, Are young borrowers bad borrowers? evidence from the credit card act of 2009, *Working Paper*.

DiPasquale, Denise, and Edward L. Glaeser, 1999, Incentives and social capital: Are homeowners better citizens?, *Journal of Urban Economics* 45, 354–384.

Fisher, Jonas D.M., and Martin Gervais, 2011, Why has home ownership fallen among the young?, *International Economic Review* 52, 883–912.

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan, 1993, Earnings losses of displaced workers, *American Economic Review* 83, 685–709.

Jiang, Wei, Ashlyn A. Nelson, and Edward Vytlacil, Forthcoming, Liar's loan? effects of origination channel and information falsification on mortgage delinquency, *Review of Economics and Statistics*.

Karlan, Dean, and Jonathan Zinman, 2010, Expanding credit access: Using randomized supply decisions to estimate the impacts, *Review of Financial Studies* 23, 433–464.

Keys, Benjamin J., Tanmoy K. Mukherjee, Amit Seru, and Vikrant Vig, 2010, Did securitization lead to lax screening? evidence from subprime loans, *Quarterly Journal of Economics* 125, 307–362.

Landvoigt, Tim, 2011, Housing demand during the boom: The role of expectations and credit constraints, *Working Paper*.

Malmendier, Ulrike, and Stefan Nagel, 2011, Depression babies: Do macroeconomic experiences affect risk-taking?, *Quarterly Journal of Economics* 126, 373–416.

——— , 2013, Learning from inflation experience, *Working Paper*.

Mian, Atif, and Amir Sufi, 2011, House prices, home equity-based borrowing, and the us household leverage crisis, *American Economic Review* 101, 2132–2156.

Morse, Adair, 2011, Payday lenders: Heroes or villains?, *Journal of Financial Economics* 102.

Ortalo-Magne, Francois, and Sven Rady, 2006, Housing market dynamics: On the contribution of income shocks and credit constraints, *Review of Economic Studies* 73, 459–485.

Saiz, Albert, 2010, The geographic determinants of housing supply, *Quarterly Journal of Economics* 125, 1253–1296.

Sinai, Todd, and Nicholas S. Souleles, 2005, Owner-occupied housing as a hedge against rent risk, *Quarterly Journal of Economics* 120, 763–789.

Sullivan, James X., 2008, Borrowing during unemployment: Unsecured debt as a safety net, *Journal of Human Resources* 43, 383–412.

Figure 1: Probability of Homeownership, 1999-2012

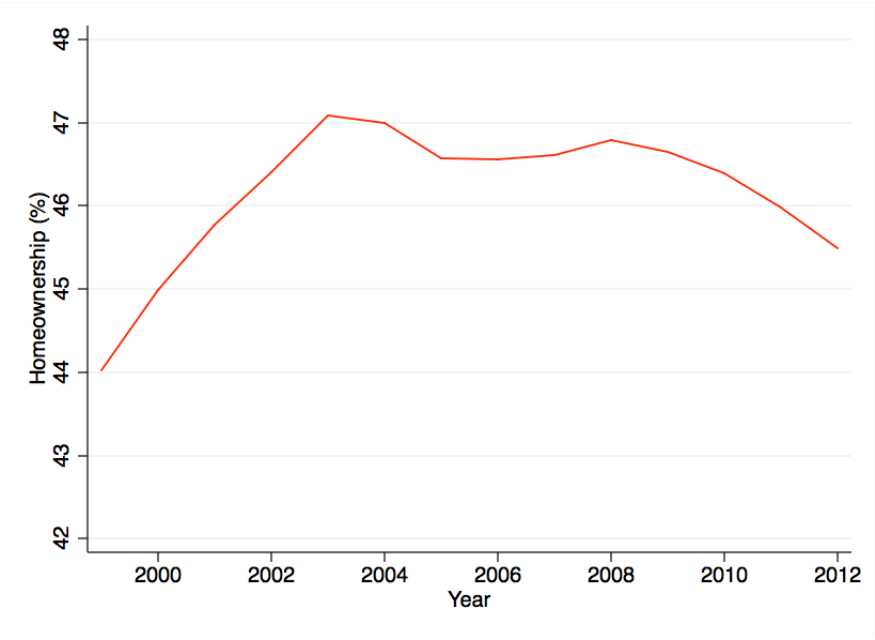
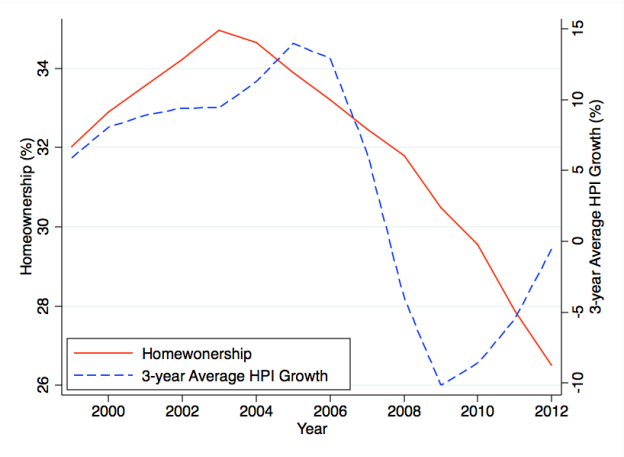
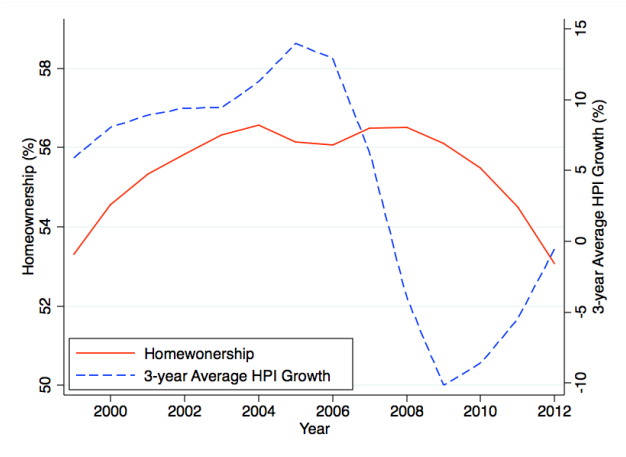


Figure 2: Probability of Homeownership by Agegroup, 1999-2012

(a) Age group: 25-34



(b) Age group: 35-44



(c) Age group: 45-60

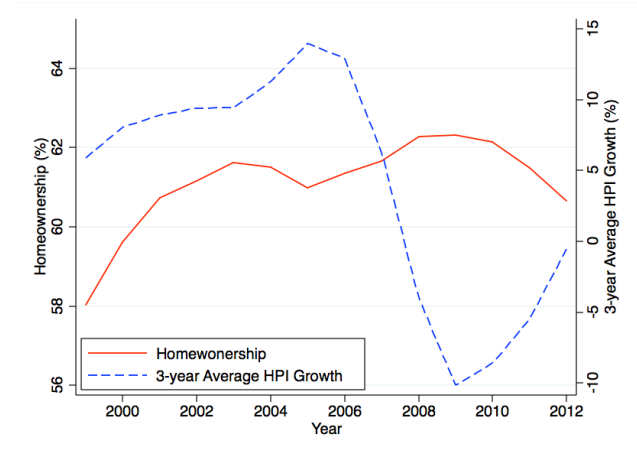


Figure 3: Hazard Rate of Buying First Home by Age

Figure 4: Cumulative Distribution of Home Ownership by Age

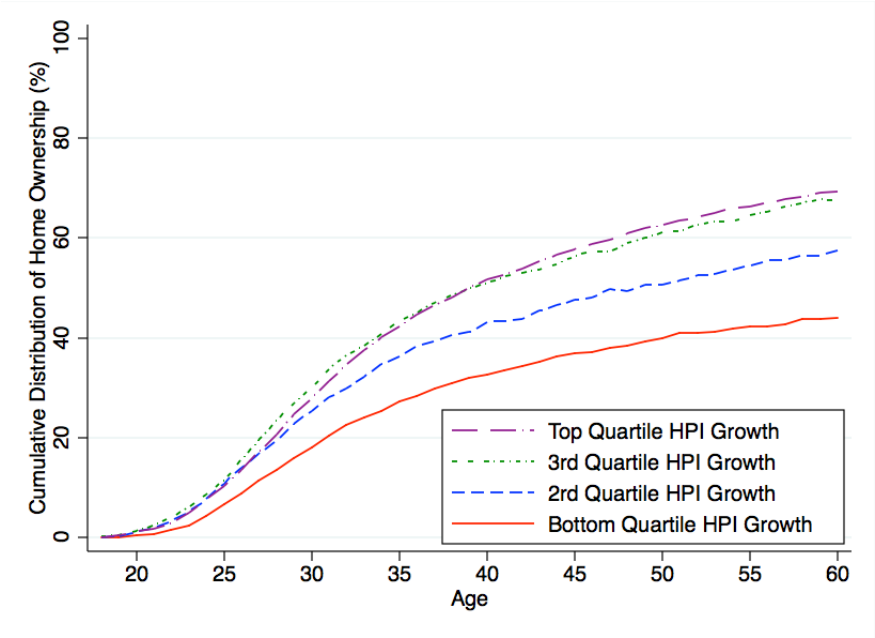
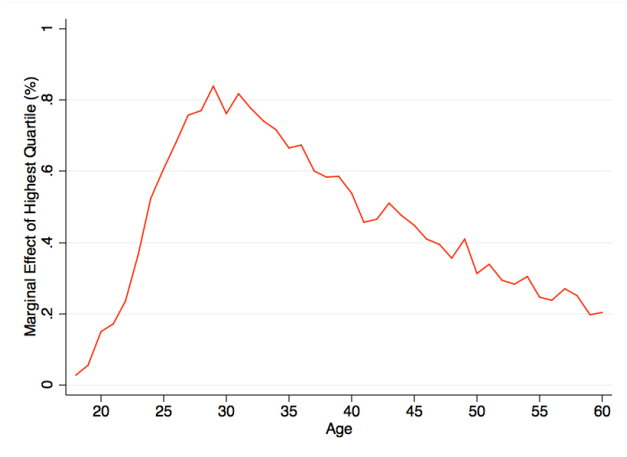


Figure 5: The HPI Effect by Age (Estimated by Logit Regression)

(a) Hazard Rate of Buying First Home by Age

(b) Marginal Effects of Highest Quartile HPI Growth



(c) Cumulative Distribution of Homeownership by Age

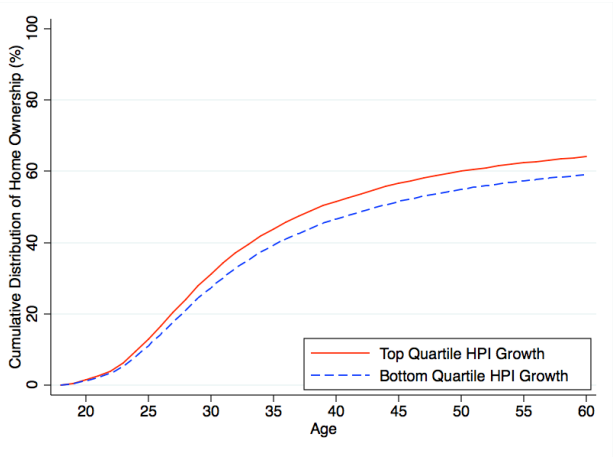
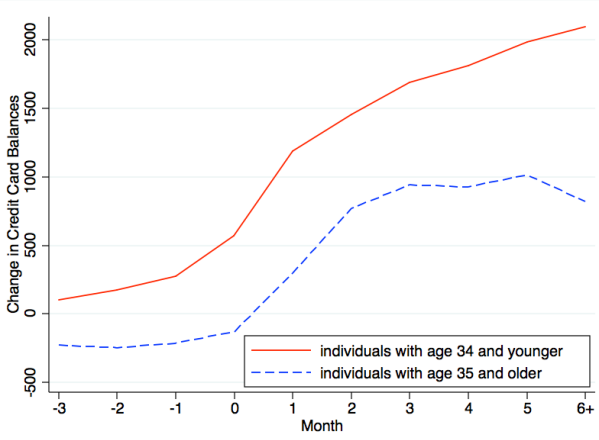
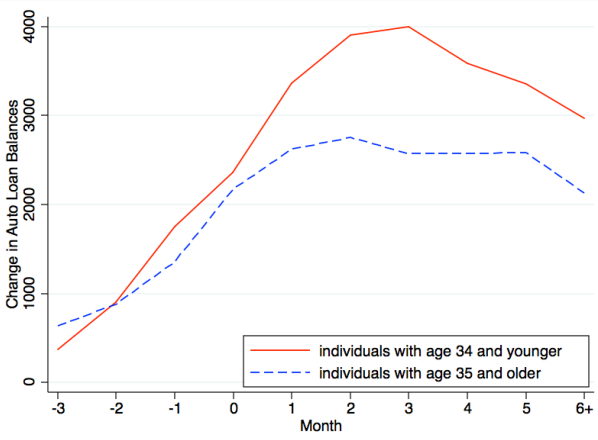


Figure 6: Change in Consumer Credits around First-Home Purchases (Younger vs Older)

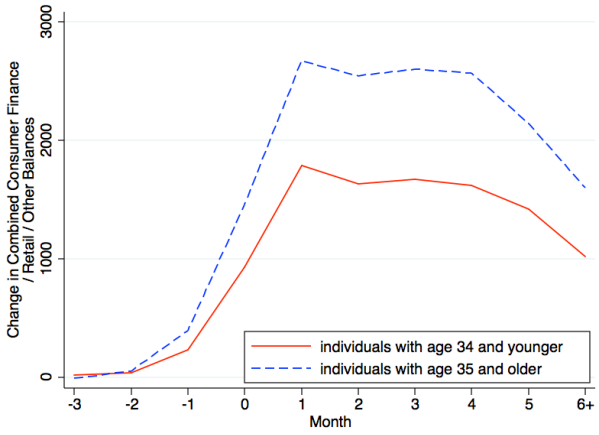
(a) Credit Card Balances
Obs = 127739
Joint test of diff significance: P-value = 0.0000



(b) Auto Loan Balances
Obs = 127739
Joint test of diff significance: P-value = 0.0000)



(c) Combined Consumer Finance/Retail/Other Balances
Obs = 1277739
Joint test of diff significance: P-value = 0.3763



(d) Student Loan Balances
Obs = 1277739
Joint test of diff significance: P-value = 0.0005

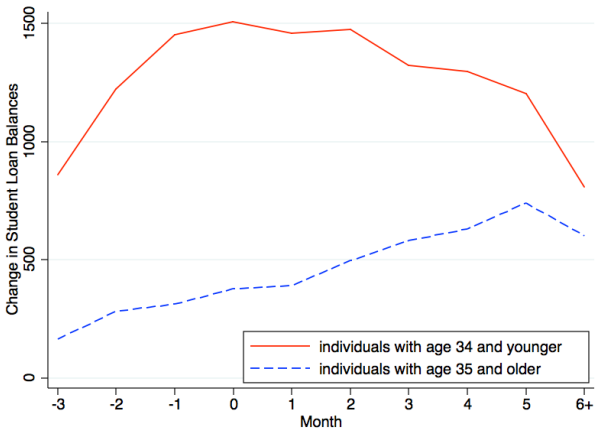
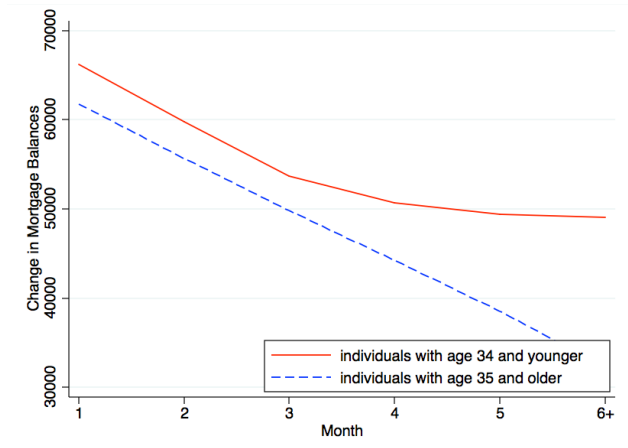
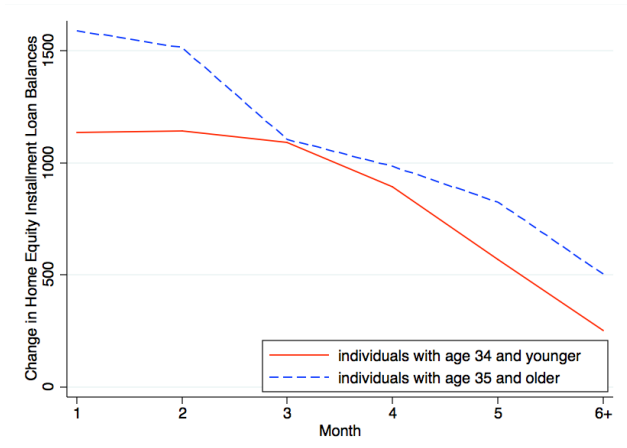


Figure 7: Change in Mortgage and Home Equity-Based Borrowing After First-Home Purchases (Younger vs Older)

(a) Mortgage Balances
Obs = 16688
Joint test of diff significance: P-value = 0.0036



(b) Home Equity Installment Loan Balances
Obs = 16688
Joint test of diff significance: P-value = 0.0245)



(c) Home Equity Line of Credit Balances
Obs = 16688
Joint test of diff significance: P-value = 0.0082

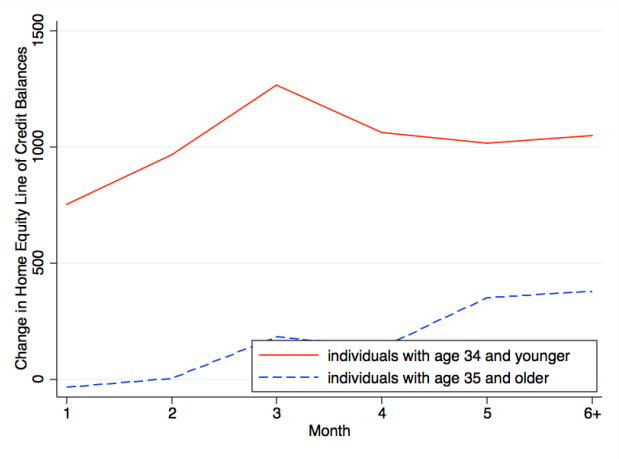


Figure 8: Change in Credit Score around First-Home Purchases (Younger vs Older)

Obs = 115994, Joint test of diff significance: P-value = 0.0000

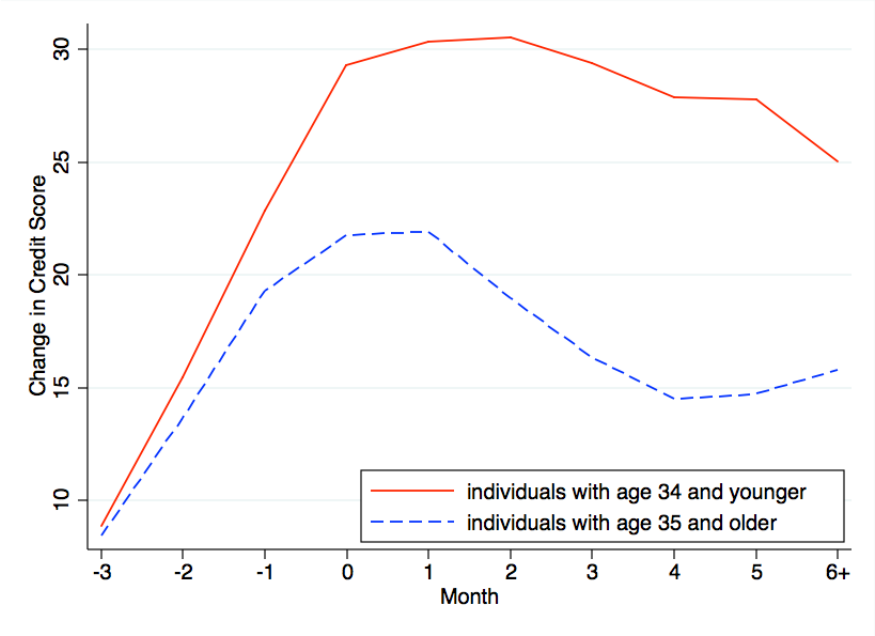
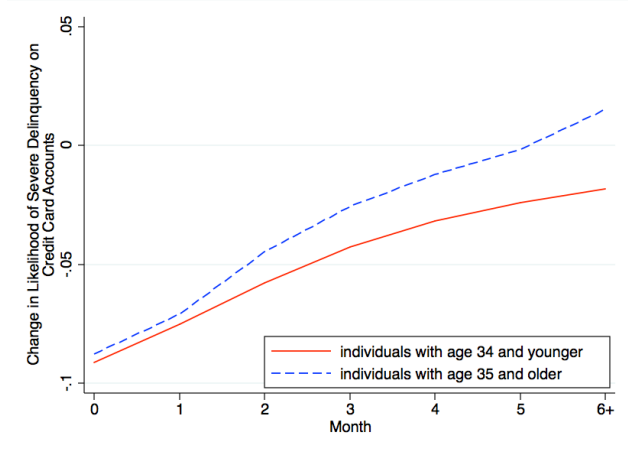
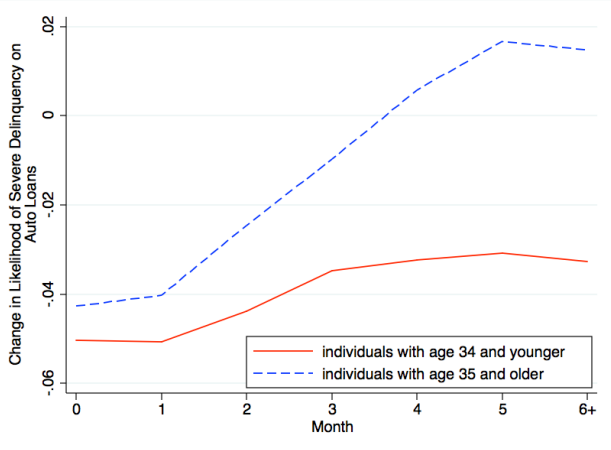


Figure 9: Change in Delinquency Rate of Consumer Credit Accounts After First-Home Purchases (Younger vs Older)

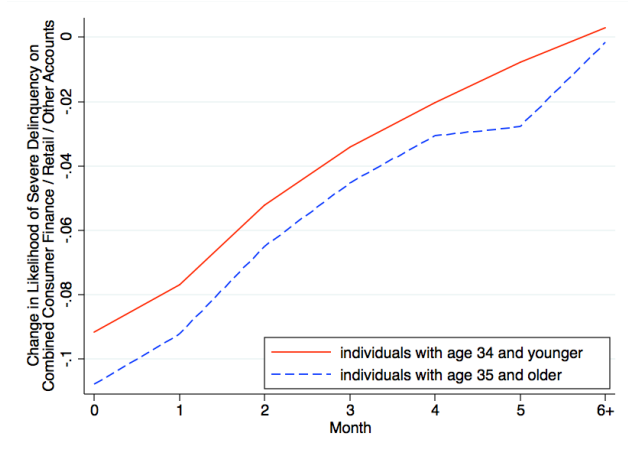
(a) Delinquency on Credit Card Accounts
Obs = 93693
Joint test of diff significance: P-value = 0.0337



(b) Delinquency on Auto Loans
Obs = 68081
Joint test of diff significance: P-value = 0.0001)



(c) Delinquency on Combined Consumer Finance/Retail/Other Accounts
Obs = 90076
Joint test of diff significance: P-value = 0.2085



(d) Delinquency on Student Loans
Obs = 24931
Joint test of diff significance: P-value = 0.0762

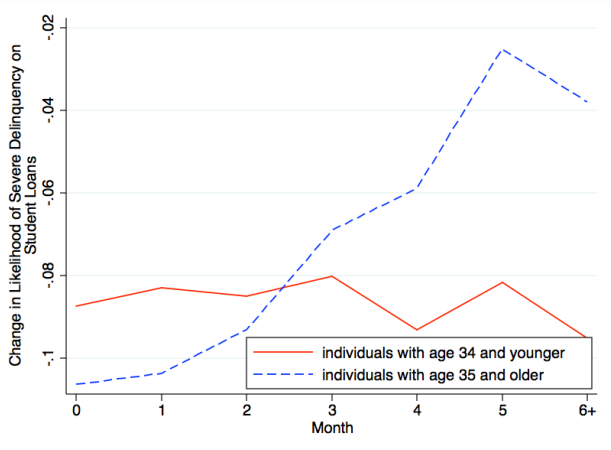
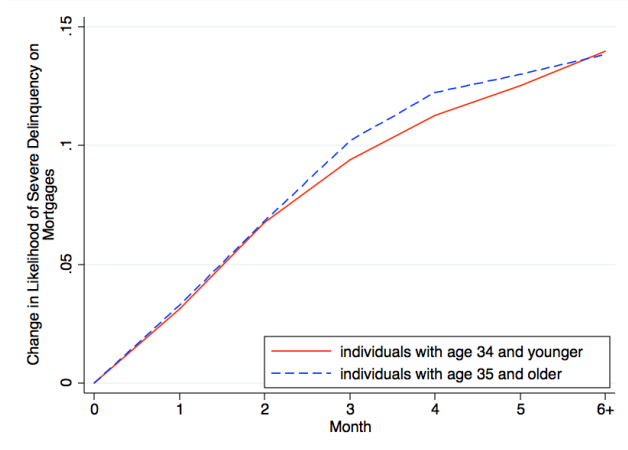
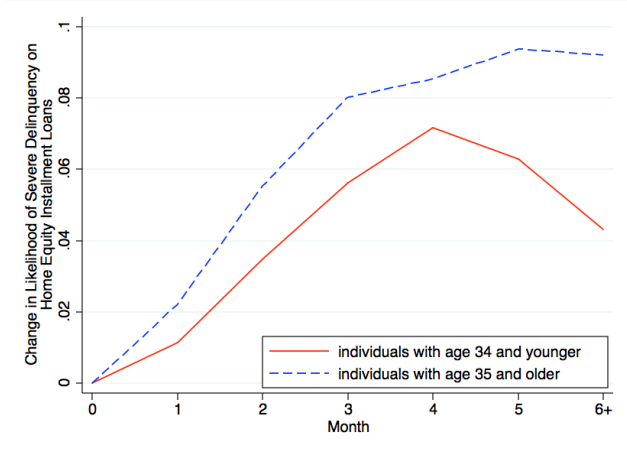


Figure 10: Change in Delinquency Rate of Mortgage and Home Equity-Based Borrowing After First-Home Purchases (Younger vs Older)

(a) Delinquency on Mortgages
Obs = 15839
Joint test of diff significance: P-value = 0.5668



(b) Delinquency on Home Equity Installment Loans
Obs = 3668
Joint test of diff significance: P-value = 0.7426)



(c) Delinquency on Home Equity Line of Credit Accounts
Obs = 3179
Joint test of diff significance: P-value = 0.5249

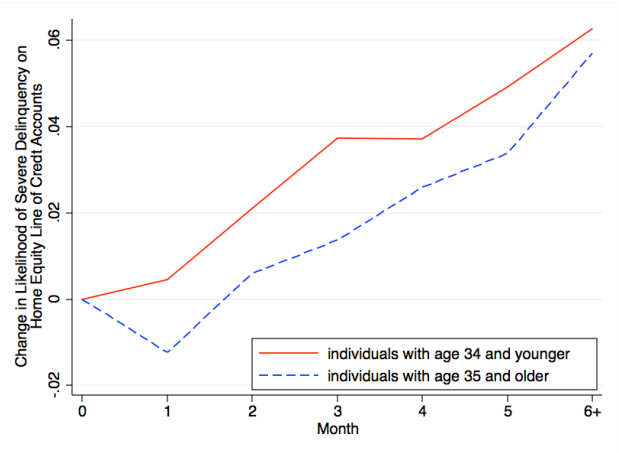


Table 1: Summary Statistics

	Obs	Mean	Std	Min	Median	Max
A. Pooled individual-year obs:						
Age	1,184,591	39.9	11.6	18	40	60
Credit Score	1,096,090	666.9	106	291	677	842
B. Individual obs:						
Age at first-home purchase	16,955	35.4	10.0	18	33	60
Credit score range for individuals over time (High-Low)	127,873	150.7	80.8	0	141	498
C. Pooled CBSA-year obs:						
3-year average HPI growth	13,379	0.029	0.06	-0.008	0.036	0.061
D. CBSA obs:						
HPI growth range for CBSA over time (High-Low)	960	0.207	0.118	0.133	0.164	0.249

Table 2: Hazard Rate of Buying the First Home (Logit Regression at the Individual Level, SE Clustered at CBSA-Level)

	(1)	(2)	(3)	(4)	(5)
HPI growth measure:	3-year HPI growth from t-3 to t			Quartile dummy for 3-year HPI growth from t-3 to t	Quartile dummy for 2-year HPI growth from t-3 to t-1
HPI growth	0.81	0.727	0.502		
	[0.244]***	[0.250]***	[0.262]*		
2nd quartile HPI growth CBSAs				0.140	0.103
				[0.043]***	[0.042]**
3rd quartile HPI growth CBSAs				0.130	0.137
				[0.049]**	[0.048]***
4th quartile HPI growth CBSAs				0.173	0.175
				[0.050]***	[0.047]***
Credit score (time-varying)	0.006	0.006	0.006	0.006	0.006
	[0.000]***	[0.000]***	[0.000]***	[0.000]***	[0.000]***
Annual average employment growth		0.007	0.004	0.004	0.004
		[0.005]	[0.006]	[0.006]	[0.006]
Average weekly wage growth		0.009	0.008	0.008	0.008
		[0.006]	[0.006]	[0.006]	[0.006]
Average quarterly growth in number of establishments		-0.002	-0.003	-0.002	-0.003
		[0.004]	[0.004]	[0.004]	[0.004]
Unemployment rate			-0.026	-0.027	-0.028
			[0.013]**	[0.013]**	[0.013]**
Observations	524,168	524,168	523,768	523,768	523,768
Year dummies	Yes	Yes	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes	Yes	Yes
CBSA dummies	Yes	Yes	Yes	Yes	Yes

Standard errors in brackets *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Appendix

A.1 Conceptual Model

A.1.1 Model Setup

Preferences. We model the timing of house purchases and consumption choices of the households who live for T periods. For the simplified case, T equals to 2. The household decides whether to purchase a house in each period. If the household buys the house at time τ, O_t equals to 1 for all $t \geq \tau$, meaning that the household owns the house. For simplicity, all the houses are in the same size. Household's choice of house size to own is not considered here. In each period t, the household also needs to optimize nondurable consumption C_t.

The household derives utility from both housing and nondurable goods for each period before period T, and from bequeathing terminal wealth, W_T. We assume owning and renting is indifferent for household at $t = 0$, but household could get a huge extra utility from owning a house rather than renting a house at $t = 1$. Therefore, to normalize the utility from renting a house as 0, the lifetime utility could be described as:

$$U_0 = logC_0 + logC_1 + logW_2 + mO_1$$

Where m is the extra utility a household could obtain if she owns a house at $t = 1$. Note that there is no extra utility if the household owns a house at $t = 0$. Let's assume m is extremely large for now, so that if households could ever afford the down payments, they will purchase the house, either at time 0 or time 1.

Housing. The household can rent or own a house to live. To simplify household's choice and emphasize the timing problem, all the houses are assumed as the same size. If the household purchases a house at period τ, $b_\tau = 1$. And for all the other periods, $b_t = 0$ $(t \neq \tau)$. The purchase price of the house at period t is denoted as P_t. The perceived future house price will affect household's decision. I'll discuss the details about perceived future house price later. If the household rents a house to live, she needs to pay a fixed rent, $\overline{N}$, for each period. For the simplified case, we normalize $\overline{N} = 0$.

Borrowing Constraint. The household could borrow against the value of the house to buy the house at a fixed rate R_D, which is normalized to zero for the simplified case. Let D_t denote the dollar amount the investor owns in mortgages at period t. Following Cocco (2005), we assume that the investor is allowed in every period to costlessly renegotiate the desired level of debt (for example, prepayment).

A down payment is required to buy a house. Specifically, the household has to pay up at least a proportion (d) of the value of the house (P_t). In other words, the mortgage value

will be less than the remaining portion of the value of the house after the down payment:

$$0 \leq D_{\{t\}} \leq (1-d)\, P_t, \quad \forall\, t$$

Beliefs of Future House Price. At this point, we want to keep it simple to focus only on the effects of borrowing constraint under the assumption of preference shift of housing in the middle age, hence we will assume the price is deterministic in order to factor out the effects of house price risk. Assume household believes that house price will increase by rate λ at $t = 1$, i.e. $P_0 = P, \quad P_1 = P_2 = (1+\lambda)\, P$, we will consider three simple cases: (**??**) stable house prices:$\lambda = 0$; (**??**) downward house prices: $\lambda < 0$; (**??**) upward house prices: $\lambda > 0$.

Labor Income. The household earns labor income for $t = 0$ and $t = 1$. To avoid the influence of labor income risk and the shape of income in the life cycle, we will assume for both periods, the household earns the same certain amount, $\frac{Y}{2}$, i.e. $Y_0 = Y_1 = \frac{Y}{2}$.

Budget Constraint. The household could save beforehand at the risk free rate, R_f. For the simplified case, let $R_f = 0$. Let S_t denote the saving of the household. The liquid wealth at period $t(t > 1)$ is $LW_t = S_{t-1} - D_{t-1}$. Following Cocco (2004), Deaton (1991) and Carroll (1997), we calculate cash-on-hand as adding period t liquid wealth to period t labor income $LW_t + Y_t$. In each period, the household needs to choose the nondurable consumption level and decide whether to buy or continue renting a house. The budget constraint at period t is given by:

$$S_t = LW_t + Y_t - C_t + D_t - b_t P_t$$

$S_t \geq 0$ *The last period wealth is given by:* $W_T = LW_T + P_T$.

Optimization Problem. The household maximizes the lifetime utility by choosing the optimal nondurable consumption ($\{C_t\}_{t=0,1}$), the optimal time to purchasing a house ($\{b_t\}_{t=0,1}$), and the optimal level of debt ($\{D_t\}_{t=0,1}$). In the simplified case, since both borrowing and saving rates are normalized to zero, households do not have preferences between savings and debts. To make it convenient for discussion, we can let the debt level always be $(1-d)\, P_t$ after the household buys a house. The next section will continue to discuss the optimal consumption choice and the best timing of buying a house under two cases.

A.1.2 Case 1: stable house price ($\lambda = 0$, i.e. $P_0 = P_1 = P_2 = P$)

In this case, households believe that house prices will maintain at the same level for all periods. There is no investment incentive to purchase a house. The benefit of buying a house only comes from the extra utility of owning a house at time 1. Since buying a house requires down payments and reduces the current consumption, households will not have incentives

to buy houses when they are young ($t = 0$). Instead, they will plan to purchase houses in their middle age ($t = 1$).

Without any constraints, households would choose the optimal consumption levels as $C_0^{NC} = C_1^{NC} = \frac{Y}{3}$. But since households have to make down payments for house purchases, they may not achieve the unconstrained optimum. Specifically,

1. If $\frac{Y}{2} - dP \geq \frac{Y}{3}$ ($\frac{dp}{Y} \leq \frac{1}{6}$), households are affluent. Even during their young age, they could buy the house and still keep the unconstrained optimal consumptions. Therefore, they will buy the house at either time 0 or time 1 ($\tau^* = 0$ *or* $\tau^* = 1$) and choose the unconstrained optimal consumptions ($C_0^* = C_1^* = \frac{Y}{3}$).

2. If $\frac{Y}{2} - dP < \frac{Y}{3}$ and $\frac{Y}{3} - dP \geq 0$ ($\frac{1}{6} < \frac{dP}{Y} \leq \frac{1}{3}$), households can afford the home at time 0, but if they do so, they will not be able to maintain consumption level C_0^{NC}. Hence, they will choose to delay the house purchase and buy it at time 1. Under this condition, since $\frac{Y}{3} - dP \geq 0$, households will still be able to keep both C_0^{NC} and C_1^{NC}. Therefore, households will buy the house at time 1 ($\tau^* = 1$) and choose the unconstrained optimal consumptions ($C_0^* = C_1^* = \frac{Y}{3}$).

3. If $\frac{Y}{3} - dP < 0$ and $\frac{Y}{2} - dP \geq 0$ ($\frac{1}{3} < \frac{dP}{Y} \leq \frac{1}{2}$), households still can afford the home at time 0, but similarly, they will choose to buy the home at time 1 ($\tau^* = 1$). But even they buy the home at time 1, they will not be able to achieve the consumption levels of C_0^{NC} and C_1^{NC}. Instead, the optimal consumption levels will be $C_0^* = C_1^* = \frac{Y}{2} - \frac{dP}{2}$.

4. If $\frac{Y}{2} - dP < 0$ and $Y - dP \geq 0$, households can not afford the down payment at time 0 any more, but through saving money in the first period, they can still purchase the home at time 1. Therefore, the optimal time of buying a home is time 1 ($\tau^* = 1$) and $C_0^* = C_1^* = \frac{Y}{2} - \frac{dP}{2}$.

5. If $Y - dP \geq 0$, the lifetime income is too low to afford the down payment, so the household will never buy a home.

The results can be summarized in A.1 When households believe that house prices will be stable in the future, they prefer to buy the houses in their middle age.

[Insert Figure A.1]

A.1.3 Case 2: downward house price ($\lambda < 0$)

In this case, at time 0, households expect that house prices will decrease in the next period. We can easily show that buying homes at time 1 is always preferred to buying homes at time 0.

Let $C^*_{0|\tau=0}$, $C^*_{1|\tau=0}$, $W^*_{2|\tau=0}$ denote the optimal consumption choices given the households buy homes at time 0. Then these choices satisfy the conditions:

$$\begin{cases} C^*_{0|\tau=0} < \frac{Y}{2} - dP \\ C^*_{0|\tau=0} + C^*_{1|\tau=0} < Y - dP \\ W^*_{2|\tau=0} = Y + \lambda P - C^*_{0|\tau=0} - C^*_{1|\tau=0} \end{cases}$$

Consider an alternative consumption bundle with buying homes at time 1 ($C_{0|\tau=1}$, $C_{1|\tau=1}$, $W_{2|\tau=1}$). If we increase the consumption of time 0 by a small positive amount, (i.e. $C_{0|\tau=1} = C^*_{0|\tau=0} + \epsilon$, $0 < \epsilon < min(dP,\ -d\lambda P,\ -\lambda P)$), and keep the consumption of time 1 the same (i.e. $C_{1|\tau=1} = C^*_{1|\tau=0}$), we can show that this alternative bundle is feasible, because

$$\begin{cases} C_{0|\tau=1} < \frac{Y}{2} \\ C_{0|\tau=1} + C_{1|\tau=1} < Y - d(1+\lambda)P \end{cases}$$

and strongly preferred to ($C^*_{0|\tau=0}$, $C^*_{1|\tau=0}$, $W^*_{2|\tau=0}$), because

$$\begin{cases} C_{0|\tau=1} > C^*_{0|\tau=0} \\ C_{1|\tau=1} = C^*_{1|\tau=1} \\ W_{2|\tau=0} = Y - C_{0|\tau=1} - C_{1|\tau=1} > W^*_{2|\tau=0} \end{cases}$$

Therefore, when individuals expect that house prices will go down in the future, they will always prefer to delay home purchases to their middle age.

A.1.4 Case 3: upward house price (λ>0)

In this case, at time 0, households expect that house prices will increase in the next period. They may consider buying homes during the young age, because (1) now buying a home at time 0 could bring up investment opportunities (investment motives of housing); (2) some of them may not be able to afford the high price at time 1 if not buying at time 0 (consumption motives of housing). We will discuss households' decisions in different regions of $\frac{\lambda P}{Y}$ and $\frac{dP}{Y}$ in more details. Then, to separate the young-age home buying induced by consumption motives, we will compare the results with the decisions under a baseline model without middle-age extra utility of owning a house.

[Insert Figure A.2]

1. If $\frac{Y}{2} - dP < 0$ and $Y - d(1+\lambda)P < 0$ (Area A in Figure A.2), households can not afford in either period. So they never buy houses.

2. If $\frac{Y}{2} - dP < 0$ and $Y - d(1+\lambda)P \geq 0$ (Area B in Figure A.2), households can not

afford at time 0 but can afford at time 1. So they buy houses at time 1.

3. If $\frac{Y}{2} - dP \geq 0$ and $Y - d(1+\lambda)P < 0$ (Area C in Figure A.2), households can afford at time 0, but will not be able to afford if they wait until time 1 to buy houses. So they buy houses at time 0.

4. If $\frac{Y}{2} - dP \geq 0$ and $Y - d(1+\lambda)P \geq 0$ (Area D1-D5 in Figure A.2), households can afford in both periods. We proceed as follows: (1) solve the optimal consumption plan given the household chooses to buy a home at time 0 and at time 1, respectively; (2) compare the utilities based on the two conditional optimized consumption plans, and decide which period is the best time to buy a home.

   Given the household buy a home at time 0, the unconstrained optimal consumption plan is: $C^{NC}_{0|\tau=0} = C^{NC}_{1|\tau=0} = \frac{1}{3}(Y + \lambda P)$.

   (a) If $\frac{Y}{2} - dP \geq \frac{1}{3}(Y + \lambda P)$ (Area D1 in Figure A.2), households can maintain the unconstrained optimal consumption and at the same time buy houses at time 0. So their best choices are to buy houses at time 0 ($\tau^* = 0$) and consume $C^*_0 = C^*_1 = \frac{1}{3}(Y + \lambda P)$.

   (b) If $0 \leq \frac{Y}{2} - dP < \frac{1}{3}(Y + \lambda P)$ (Area D2-D5 in Figure A.2), households are not able to maintain $C^{NC}_{0|\tau=0}$ if buying houses at time 0. So the corner solution gives that, at time 0, households consume $C^*_{0|\tau=0} = \frac{Y}{2} - dP$. To further split the region,

   - If $0 \leq \frac{Y}{2} - dP < \lambda P$ (Area D2 and D3 in Figure A.2), which means $\frac{Y}{2} < (d+\lambda)P$, the income at time 1 is less than the final wealth after selling the house at time 2. Since the households cannot borrow from the future other than buying the house in the simplified model, they are not able to smooth consumption between time 1 and time 2. So they consume $C^*_{1|\tau=0} = \frac{Y}{2}$, hence $W^*_{2|\tau=0} = (d+\lambda)P$. Let's denote this non-smoothing consumption bundle as $\mathbb{C}^{NS}|_{\tau=0} = \{\frac{Y}{2} - dP, \ \frac{Y}{2}, \ (d+\lambda)P\}$.
   - If $\lambda P \leq \frac{Y}{2} - dP < \frac{1}{3}(Y + \lambda P)$ (Area D4 and D5 in Figure A.2), households will smooth the consumption between time 1 and time 2, so $C^*_{1|\tau=0} = W^*_{2|\tau=0} = \frac{1}{2}\left(\frac{Y}{2} + (d+\lambda)P\right)$. Let's denote this smoothing consumption bundle as $\mathbb{C}^{S}|_{\tau=0} = \{\frac{Y}{2} - dP, \ \frac{1}{2}\left(\frac{Y}{2} + (d+\lambda)P\right), \ \frac{1}{2}\left(\frac{Y}{2} + (d+\lambda)P\right)\}$.

   Given the household buy a home at time 1, the unconstrained optimal consumption plan is: $C^{NC}_{0|\tau=1} = C^{NC}_{1|\tau=1} = \frac{Y}{3}$.

(a) If $\frac{Y}{3} \geq d(1+\lambda)P$ (Area D2 and D4 in Figure A.2), even though households buy the houses at a high price, they can still maintain the unconstrained consumption level $\frac{Y}{3}$. Let's denote this smoothing consumption bundle as $\mathbb{C}^S|_{\tau=1} = \{\frac{Y}{3}, \frac{Y}{3}, \frac{Y}{3}\}$.

(b) If $\frac{Y}{3} < d(1+\lambda)P$ (Area D3 and D5 in Figure A.2), households can not achieve the unconstrained consumption level if they buy houses at time 1. So the corner solution gives the non-smoothing consumption bundle as $\mathbb{C}^{NS}|_{\tau=1} = \{\frac{1}{2}(Y - d(1+\lambda)P), \frac{1}{2}(Y - d(1+\lambda)P),\ d(1+\lambda)P\}$.

We then compare the conditional maximized utilities given buying homes at time 0 and buying homes at time 1 in Area D2-D5. Through comparison, we can obtain the regions of optimal timing of house purchases as shown in A.3. The horizontal line represents the special case of stable house prices (Case 1). The blue lines separate the area into three regions. For notation convenience, we denote $L = \frac{\lambda P}{Y}$ and $K = \frac{dP}{Y}$. The region (R) where households choose to buy houses at time 0 could be described as:

$$\begin{cases} L \geq 0 & \text{when } K \leq \frac{1}{6} \\ L \geq S(K), \text{where } S(\frac{1}{6}) = 0,\ S(\frac{1}{2}) = \frac{1}{2d},\ S^{'}(K) > 0,\ S^{''}(K) > 0 & \text{when } \frac{1}{6} < K \leq \frac{1}{2} \\ L > \frac{1}{2d} & \text{when } K = \frac{1}{2} \end{cases}$$

When price expectation is higher than the threshold, households buy homes at their young age (time 0). They are willing to sacrifice their consumption in the young age to either profit from the investment opportunity or hedge to assure their home ownership during their middle age.

A.1.5 Testable Prediction

We assume the population of households is heterogeneous in their lifetime income Y with c.d.f. F. The following proposition outlines the testable prediction regarding the probability of buying homes at time 0 ($Pr(\tau^* = 0)$).

PROPOSITION 1. *All else equal, the probability of buying homes at time 0,* $\Pr(\tau^* = 0)$*, increases with price expectation parameter* λ*, decreases with the current price* P*.*

Sketch of Proof: For $\lambda < 0$, since households always prefer to buy houses at time 1, so $Pr(\tau^* = 0) = 0$;

For $\lambda \geq 0$, the probability of buying homes at time 0 is: $Pr(\tau^* = 0) = \int_{(L,K)\in R} dF$, which is greater than zero. We discuss two cases $0 \leq \lambda < 1$ and $\lambda \geq 1$ respectively. Let's write the relationship between L and K as $L(K) = \frac{\lambda}{d}K$.

If $\lambda < 1$, then $L(\frac{1}{2}) < S(\frac{1}{2})$.We also know that $L(\frac{1}{6}) > S(\frac{1}{6})$. To compare $L(K)$ and $S(K)$ for $\frac{1}{6} < K \leq \frac{1}{2}$, we look at their difference $F(K) = L(K) - S(K)$. So we have $F(\frac{1}{2}) > 0$

and $F(\frac{1}{6}) < 0$. We can easily get that there is a unique solution$K^c \in (\frac{1}{6}, \frac{1}{2}]$ s.t. $F(K^c) = 0$ and $F^{'}(K^c) < 0$. It means that, when $\frac{1}{6} < K < K^c$ (i.e. $Y^c < Y < 6dP$, where $\frac{dP}{Y^c} = K^c$), households choose to buy homes at time 0. At the same time, we know, for $0 < K \leq \frac{1}{6}$, $L = \frac{\lambda}{d}K \geq 0$ for any Y $s.t.$ $0 < K \leq \frac{1}{6}$ (i.e. $Y \geq 6dP$).

Therefore, we can get that $Pr(\tau^* = 0) = \int_{Y^c}^{\infty} dF$. We can further proof that $\frac{\partial Y^c}{\partial \lambda} = -\frac{\frac{P}{Y^c}}{F'(K^c)\cdot(-\frac{dP}{{Y^c}^2})} < 0$. Therefore, $\frac{\partial \Pr(\tau^*=0)}{\partial \lambda} > 0$.

If $\lambda \geq 1$,then $L(\frac{1}{2}) \geq S(\frac{1}{2})$. Given$L(\frac{1}{6}) > S(\frac{1}{6})$, then for the whole region $K \in (0, \frac{1}{2}]$, households choose to buy homes at time 0. So, $\Pr(\tau^* = 0) = \int_{2\mathrm{dP}}^{\infty} dF$ and$\frac{\partial \Pr(\tau^*=0)}{\partial \lambda} = 0$.

After combining the cases, we get $\frac{\partial \Pr(\tau^*=0)}{\partial \lambda} \geq 0$.

PROPOSITION 2. *All else equal, the probability of buying homes at time 0,* $\Pr(\tau^* = 0)$, *decreases with the current price* P.

Figure A.1: Regions of Opitimal Time of First-Home Purchases and Optimal Consumption Plan (Case 1)

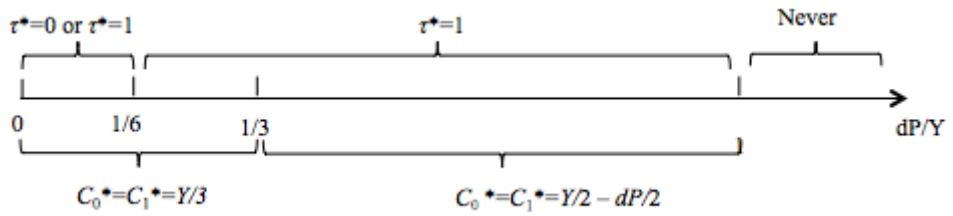


Figure A.2: Regions for Discussing Optimal Timing of First-Home Purchases (Case 3)

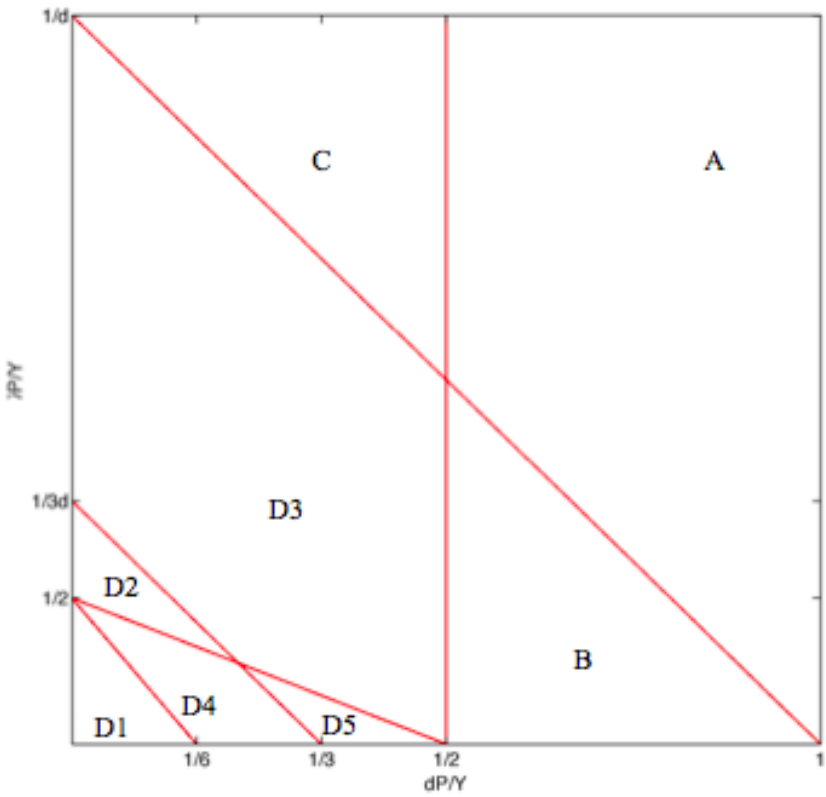
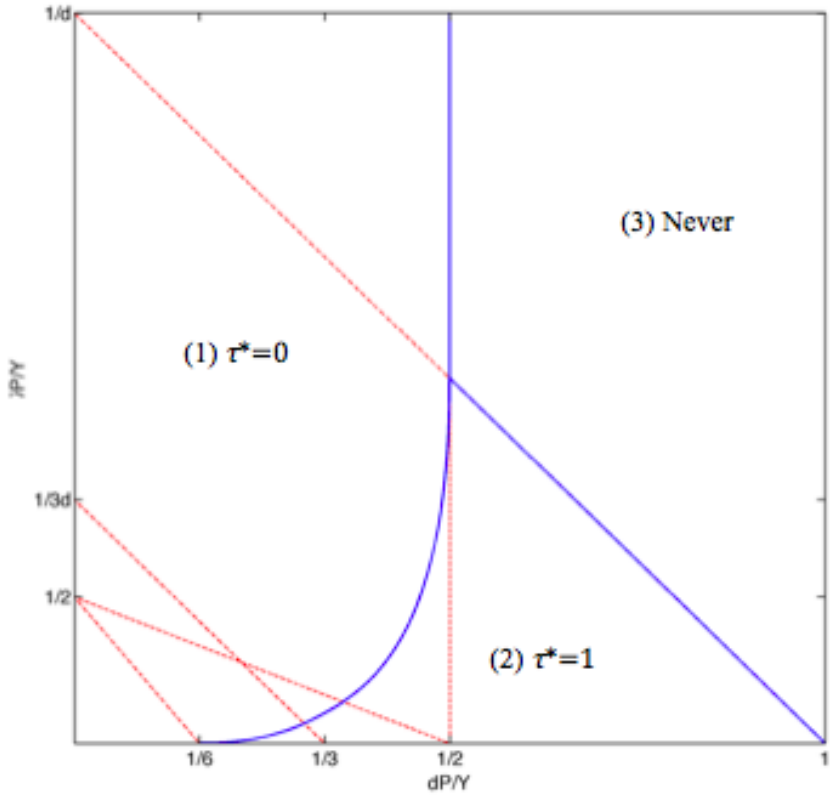


Figure A.3: Regions of Optimal Timing of First-Home Purchases (Case 3)



A.2 Robustness Tests for Hazard Rate of Buying the First Home

Table A.1: Hazard Rate of Buying the First Home (Instrumental Variable Estimation)

	(1)	(2)
	First Stage	IV Probit
Left-hand side variable:	3-year HPI growth from t-3 to t	Hazard rate
HPI growth measure:		Instrumented 3-year HPI growth
Elasticity × National HPI growth	-0.278	
	[0.032]***	
HPI growth		0.525
(Instrument = Elasticity × National HPI growth)		[0.311]*
Credit score (time-varying)	0	0.003
	[0.000]	[0.000]***
Annual average employment growth	0.003	0.004
	[0.001]**	[0.005]
Average weekly wage growth	0.001	0.004
	[0.001]	[0.004]
Average quarterly growth in number of establishments	0.001	-0.001
	[0.001]*	[0.003]
Unemployment rate	-0.021	0.004
	[0.003]***	[0.012]
Observations	295,746	295,746
Year dummies	Yes	Yes
Age dummies	Yes	Yes
CBSA dummies	Yes	Yes

Table A.2: Hazard Rate of Buying the First Home (Subsample for Individuals Who Enter the Sample Between 22 and 25 Years Old)

	(1)	(2)	(3)
	3-yr HPI growth		2-yr HPI growth
HPI growth	0.388		
	[0.180]**		
2nd quartile HPI growth CBSAs		0.086	0.063
		[0.028]***	[0.031]**
3rd quartile HPI growth CBSAs		0.111	0.088
		[0.036]***	[0.038]**
4th quartile HPI growth CBSAs		0.142	0.135
		[0.037]***	[0.040]***
Credit score (time-varying)	0.009	0.009	0.009
	[0.000]***	[0.000]***	[0.000]***
Annual average employment growth	0.002	0.002	0.003
	[0.005]	[0.005]	[0.005]
Average weekly wage growth	0.007	0.006	0.006
	[0.004]	[0.005]	[0.005]
Average quarterly growth in number of establishments	0.000	0.000	0.000
	[0.004]	[0.005]	[0.004]
Unemployment rate	-0.024	-0.025	-0.025
	[0.011]**	[0.010]**	[0.010]**
Observations	853,285	853,285	853,285
Year dummies	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes
CBSA dummies	Yes	Yes	Yes