Federal Reserve Bank of Chicago

Rushing into the American Dream? House Prices Growth and the Timing of Homeownership

Sumit Agarwal, Luojia Hu, and Xing Huang

REVISED
May 2016

WP 2013-13

Rushing into the American Dream? House Prices Growth and the Timing of Homeownership*

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

Abstract

We use the New York Fed Consumer Credit Panel dataset to empirically examine how past house price growth influences the timing of homeownership. We find that the median individual in metropolitan areas with the highest quartile house price growth becomes a homeowner 5 years earlier than that in areas with the lowest quartile house price growth. The result is consistent with a life-cycle housing-demand model in which high past price growth increases expectations of future price growth thus accelerating home purchases at young ages. We show that extrapolative expectations formed by home-buyers are a necessary channel to explain the result.

Keywords: Housing, Homeownership, Consumer Finance, Credit Constraints, Life Cycle

JEL Classification: R21, D12, D91, D14

*The authors would like to thank Anna Neumann and Ashley Wong for excellent research assistance. We also benefited from conversations and comments from Zahi Ben-David, Souphala Chomsisengphet, Bo Honore, Amit Seru, seminar participants and discussants at Chicago Fed, Dallas Fed Housing, Stability and the Macroeconomy conference, AAEA, and IBEFA. 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 often said to be an integral part of the American dream. However, modeling the demand for housing is complicated. Houses provide utility and serve as collateral for additional credit needs. They may bring investment benefits as well. These features suggest that homeownership may vary over the life cycle and among different cohorts. Past literatures have focused on the impact of demographics (such as marriage), income, and credit constraints on homeownership. However, these factors cannot explain the sharp rise in homeownership leading up to the crisis.

In this paper, we investigate the influence of house price growth on the demand for homeownership over the lifecycle.[1] We focus on first home purchases, because they account for 40% of home sales over the past 30 years and more than 50% in 2009, according to the National Association of Realtors. First home purchases may also matter for the long-term dynamics of the housing market, because first home purchases affect the demand for trade-up homes in the future housing market.

In a life-cycle model, there may be two offsetting effects of house price growth on housing demand. On the one hand, if the 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 down payment if she wants to buy it when she is young. When she is financially constrained and her preference for owning a home varies over the life cycle, she will prefer to postpone the home purchase until middle age. Hence, house price growth decreases the probability of individuals buying first homes at early ages, which we call the *liquidity constraint channel hypothesis*. On the other hand, individuals may extrapolate future house price growth from past house price growth. These people are more likely to buy a house early in their life cycle if they expect the house prices to rise fast.[2] In this case, house price growth increases the probability of individuals buying their first homes earlier through investment and/or hedging incentives, which we call the *expectation channel hypothesis*.

In this paper, we lay out a simple conceptual model to illustrate the two different channels through which housing price could affect demand for homeownership over the life cycle. We then empirically test which effects dominate in the data. We exploit a large panel of individuals from the Federal Reserve Bank of New York's Consumer Credit Panel (CCP) data set. We follow each individual from 1999 to 2012 and study the homeownership timing decision as a function of house price changes across CBSAs controlling for other demand and supply factors. This data set is truly unique and well suited for our line of inquiry because of the following three reasons: (i) The data are a long panel of quarterly information about limited demographics

[1] Other related papers studying the impact of house prices have focused on the feedback effect on current home owner's home equity-based borrowing (Mian and Sufi, 2011), the impact on local demand and retail prices (Stroebel and Vavra, 2015), or the timing of selling (Qian, 2013).

[2] Survey evidence from the *Michigan Survey of Consumers* also suggests that a significant portion of respondents, around 30%, explicitly mention house prices to justify their views on whether the present is a "good time to buy [housing] for investment" (Piazzesi and Schneider, 2009).

and risk measures that are time varying with little measurement errors (unlike the survey data sets that can be potentially biased, e.g., the *Survey of Consumer Finance*); (ii) the data have detailed geographic information up to the zip code level; and (iii) unlike survey data sets, this is a panel of the entire population that is active in the credit markets, 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 data set, 13 years is not sufficient to study life-cycle 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 econometricians, we face a right censoring problem. To explicitly account for 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 potential home buyers (as some will never buy a home).

Our main variable of interest is the past three-year house price changes at the metropolitan area (CBSA) level. There could be confounding factors that could explain the timing of the homeownership decision. To deal with this, we have CBSA-fixed effects to control for the level of house prices across CBSAs and the credit supply that is not time varying. These effects also control 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 CBSA-specific time-varying control variables such as the unemployment rate, the growth rate of the number of businesses, and wages. We also control for individual-specific time-varying credit risk by each person's credit risk score (this measures the individual's ability and willingness to pay credit).

To fix ideas, let us explain the thought experiment that we have in mind. We take two random people and assign them to the highest and lowest house price growth CBSA quartiles and assume that they do not move between CBSAs. Then we study if the house price growth in the quartiles affects the timing of their first home purchase. We show that the probability of buying a house is accelerated in CBSAs in the highest quartile of house price growth relative to that in CBSAs with the lowest quartile of house price growth. Specifically, we find that the age at which the median individual becomes a homeowner is 5 years younger for people in the highest house price growth quartile compared with those in the lowest house price growth quartile. This magnitude is similar to what is found in Fetter (2013) — the median individual becomes a home owner about 5 years earlier if the Veteran Affairs (VA) benefits were extended to all individuals (about a 10 percentage point reduction in down payment).

One might still be concerned about the potential endogeneity of housing price growth if the controls in the model do not capture all the unobservables that affect the local housing demand through changes in fundamentals (such as demographics or economic conditions). To partially address this issue, we also instrument local house prices with national house prices interacted with local house supply elasticity (Saiz, 2010), and we still find the accelerating effect of house price growth.

The results show the effects of the expectation channel dominate. According to a survey of U.S. home buyers from 2003 through 2012 conducted by Case et al. (2012), home buyers' expectations of high future home prices are substantially affected by recent experience. We also provide supportive evidence that for the four counties documented in the survey, the expectations are positively correlated with the house price growth measure used in our paper. In the region with high house price appreciation, individuals are more likely to form expectations of high future house prices and become homeowners at young ages, despite their needing to sacrifice consumption to afford large down payment.

The extrapolative expectations may apply not only to home buyers but also to credit suppliers, who could have relaxed buyers' borrowing constraints by lowering lending standards or payments if they also have expectations of high future price growth. The relaxation of borrowing constraints may be confounded with the accelerating effect of extrapolative expectations from home buyers. To separate these two effects (mortgage-supply explanations vs house-demand explanations), we first extend our sample back to the 1990s when the mortgage supply varied little across different regions and find that the accelerating effect still exists. Then, we repeat our main analysis on the non-subprime borrowers, a group that was less likely to be much affected by the credit supply change during the recent housing cycle. Specifically, we look at a subsample of borrowers with credit scores above 660 (Keys et al., 2010) when they enter our sample. The results are largely unchanged. Furthermore, using a new merged data set between the Equifax and the Loan Performance Services (LPS), we examine how house size and down payment vary with past house price growth, and compare the changes between young and old buyers. We find that in areas with faster house price growth, young buyers tend to purchase smaller houses and, for houses of same size, they also pay higher (rather than lower) down payment. The evidence on these two dimensions can not be explained if the mortgage-supply explanation were the only force for the accelerating effect. The house-demand channel is a necessary mechanism for one to explain all evidence documented in the paper. We want to emphasize that the mortgage supply-side explanation and the housing demand-side explanation are not mutually exclusive, and we view the house-demand side story as complementary, rather than alternative, to the mortgage-supply side explanations.

Our paper is most closely related to Landvoigt (2011), in the sense that he also studies the role of credit constraints and house price expectations on the decision to purchase and the size of home purchased. 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 of 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 data set 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 housing demand over the life cycle. Ortalo-Magne and Rady (2006) study a life-cycle model of the housing market with a property ladder and a credit constraint. Rather than emphasizing on the link between income shock and house prices like they do, our paper focuses on the influence of house price expectations on the behavior of first time home buyers. Attanasio et al. (2012) also construct a life-cycle model and incorporate some realistic features, but this model still lacks one feature considered in our paper: the possibility that individuals extract time-varying utility from homeownership. Sinai and Souleles (2005) model the demand for homeownership as the trade-off between rent risk and the asset price risk. They relate the demand for homeownership to local rent volatility and individuals' expected horizon. Han (2010) identifies 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 purchases by marginal first time home buyers.

The paper proceeds as follows. In Section 2, we lay out a conceptual model highlighting the different channels through which housing prices could affect demand for homeownership over the life cycle. In Section 3, we describe the data sets we use and present some summary statistics. In Section 4, we empirically examine the impact of house price growth on the timing of first home purchases. We discuss and test some alternative explanations (such as expansion of mortgage supply) in Section 5. And Section 6 concludes.

2. A Conceptual Model

To illustrate how housing prices could affect individuals' demand for homeownership over the life cycle, we consider a simple model in which an individual maximizes her lifetime utility by choosing an optimal path for nondurable consumption, an optimal time to purchase a house, and an optimal level of debt/saving.

2.1 MODEL SETUP

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

Individuals derive utility from both housing and nondurable goods for each period before period T, as well as from bequeathing terminal wealth, W_T. We assume individuals are indifferent to owning versus renting a house at $t = 0$, but individuals 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 an individual could obtain if she owns a house at $t = 1$. Note that there is no extra utility if she owns a house at $t = 0$. Let's assume m is extremely large, so that if she could ever afford the down payment, she will purchase a house, either at time 0 or time 1.

Housing. The individual can rent or own a house in which to live. To simplify an individual's choice and emphasize the timing problem, all the houses are assumed to be the same size. If the individual 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 the individual's decision. We'll discuss the details about perceived future house price later. If the individual 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 individual 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 (2004), we assume that the investor is allowed in every period to costlessly renegotiate the desired level of debt (for example, with penalty-free prepayment).

A down payment is required to buy a house. Specifically, the individual 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 things simple to focus only on the effects of the borrowing constraint under the assumption of a preference shift of housing in middle age; hence, we will assume the price is deterministic in order to factor out the effects of house price risk. Assuming individuals believe that house prices will increase by rate λ at $t = 1$, i.e., $P_0 = P,\ P_1 = P_2 = (1+\lambda)P$, we will consider three simple cases: (1) stable house prices ($\lambda = 0$); (2) downward house prices ($\lambda < 0$); (3) upward house prices ($\lambda > 0$).

Labor Income. The individual 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 individual earns the same certain amount, $\frac{Y}{2}$, i.e., $Y_0 = Y_1 = \frac{Y}{2}$.

Budget Constraint. The individual 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 by adding period t liquid wealth to period t labor income $LW_t + Y_t$. In each period, the individual 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

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

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

Optimization Problem. The individual maximizes lifetime utility by choosing the optimal nondurable consumption ($\{C_t\}_{t=0,1}$), the optimal time for 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 individual buys a house. In Appendix A, we discuss in detail the optimal consumption choice and the best timing for buying a house under three cases.

2.2 PREDICTIONS

As described before, we assume individuals' preferences for owning a house vary over the life cycle. An individual could obtain huge extra consumption utilities from homeownership at middle age relative to at a young age. In this model, house price growth could affect the housing demand through two channels.

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

Proof. In Appendix A. Q.E.D.

The first channel is *house price expectations.* Specifically, individuals are more likely to buy a house early in their life cycle if they expect house prices to rise faster. The intuition is as follows. When there are expectations of low future house prices, the individual would prefer to postpone her purchase of a first home. Otherwise, she has to sacrifice her consumptions at a young age for the down payment without experiencing as much utility from homeownership as in middle age. In contrast, when there are expectations of high future house prices, the individual would make the first home purchase earlier. There are two incentives to do so. The first incentive is investing. The individual might obtain potential capital gain from the house purchase. The second incentive is hedging. The individual can hedge against the possibility that she may not be able to afford the down payment in middle age (thus, losing out on the extra utility from homeownership) if the price reaches sky high. Through this channel, house price growth increases the probability of individuals buying a first home earlier.

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

One caveat is that our model only illustrates the decision-making of individuals and abstracts away from the action of mortgage credit suppliers. Specifically, we assume a fixed minimum down payment ratio, which may vary with market conditions in a general equilibrium setting. If instead we assume mortgage credit suppliers would relax borrowing constraints given expectations of high future house prices, there will be two scenarios: (1) if home buyers have negative or zero price growth expectations, they will prefer to wait to buy at time 1 or be indifferent about when the purchase is made (at time 0 or time 1). In this case, relaxing borrowing constraints does not influence the timing of individual's first home purchases; (2) if home buyers share similar positive expectations of house price growth as the credit suppliers, relaxing borrowing constraints will then amplify the accelerating effect of homebuyer price growth expectations on first home purchases. Later in the paper, we'll test both mortgage-supply explanation and housing-demand explanation for the accelerating effect of housing price growth expectations on first home purchases. These two explanations are not mutually exclusive. We find that mortgage-supply alone could not explain the accelerating effect, and the housing demand explanation plays an important role.[3]

3. Data

3.1 FRBNY (EQUIFAX) CONSUMER CREDIT PANEL

The main data set we use for the empirical analysis is from the Federal Reserve Bank of New York's (FRBNY) Consumer Credit Panel (CCP). This is a panel data set collected by the credit bureau (Equifax) each quarter starting from the first quarter of 1999. Individuals in the panel are selected randomly from the U.S. population based on the last two digits of their Social Security number (SSN). 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 SSN. This means that as soon as a young adult with a randomly selected SSN opens the first line of credit

[3] Details are illustrated in Appendix A.

(often around age 18), that person will be added to the Equifax data set. Deceased individuals are dropped from the data set. This sampling methodology ensures that the 5% random sample reflects the current demography of the U.S. population with a credit history and SSN.[4]

In this paper, we use the 1% sample of the primary-individual data over the period 1999-2012 and only include individuals aged 18-60. Credit files for the very young and the very old are often small or incomplete. In particular, elderly individuals may have already paid off their mortgage in the past so that it no longer appears in their credit file,[5] which would make it impossible to identify the time of their first home purchase. So we exclude them from our analysis.

The main variables we select from the Equifax data set are a person's age, address, credit score, and mortgage history. Panel A of Table I reports summary statistics. In the sample, after all the individual-year observations are pooled, the average age is about 40 and the average credit score is 667.[6]

For our study, the key variable of interest is each person's oldest mortgage account, which we use to infer their first home purchase. Since we want to analyze home purchase decisions at a yearly frequency, we select each person's age, credit score and address at the beginning of each year (usually Q1, unless an individual enters the data set mid-year). We then look to see if they purchased their first home at any point during that calendar year. As Panel A of Table I shows, the average age to purchase the first home is 35.4 in our sample.

We exclude observations with missing birth years, which are about 5% of individuals in the Equifax sample. To the extent that birth years are missing from the data randomly, this should not affect our analysis. For individuals with inconsistent values of the age of the oldest mortgage, we take the age of the oldest mortgage from the earliest survey response where a person had reported having ever taken out a mortgage, and then replace any subsequent oldest mortgage values, if different, to match the oldest non-missing survey response.

Besides mortgages, the Equifax data also has extensive information on non-mortgage debts and their repayment status, which we will exploit in the second part of the empirical analysis. Panel A of Table 1 also presents summary statistics for the balances and the delinquency status for different credit accounts.

---

[4] See Lee and van der Klaauw (2010) for an excellent introduction to the data set.

[5] According to the FRBNY staff report (Lee and van der Klaauw, 2010), 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, p.2, Lee and van der Klaauw (2010)). However, detailed information on specific accounts (such as a mortgage) must be updated by an individuals creditors in the past 3 months in order to be included in the Equifax dataset for a given quarter: While records will be included in our panel for all individuals with some credit activity on their credit reports over the past 7 years or so, records will only include information on recently updated accounts (footnote 14, p.10, Lee and van der Klaauw (2010)).

[6] The risk score in the Equifax data ranges between 280 and 850.

3.2 CORELOGIC HOME 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 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 metropolitan area (CBSA level) to capture variation in price growth between different metropolitan areas across the country.

Specifically, we first 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:[7]

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

We then 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 I, Panel B). There is variation both across cities and over time. We also 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 I, Panel B, within a given CBSA, HPI growth varies over time — the average range of HPI growth 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 the quarter of cities with lowest HPI growth rate, it swung by 25 percentage points or more in the quarter of cities with highest HPI growth rate.

3.3 BLS: EMPLOYMENT AND WAGE DATA

To capture local business and labor market conditions, we use regional employment and wage data from the *Quarterly Census of Employment and Wages* (QCEW), compiled by the U.S. Bureau of Labor Statistics (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 the 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 Area Unemploy-

[7] For the sake of robustness, we also exploit the growth rate from quarter 4 of year $t-1$ to quarter 4 of year t, and the results remain the same.

ment 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 for each of these four macroeconomic indicators across all counties within a given CBSA. The summary statistics are reported in Table 1, Panel C.

[Insert Table I]

3.4 LOAN PERFORMANCE SERVICES (LPS) DATASET

To further test the housing-demand from the mortgage-supply explanations, we also merge the Equifax dataset with the Loan Performance Services (LPS) dataset to obtain information about loan characteristics such as loan-to-value (LTV) ratio and downpayment, etc.

The Equifax Credit Risk Insight Servicing McDash (CRISM) database links individuals in the Equifax Consumer Credit Panel to the mortgage-level McDash servicing data (LPS) using a confidential and proprietary matching process. Using the unique consumer identifier, we find each first-time homebuyer in our original sample in the 2005-2012 CRISM dataset in order to obtain mortgage characteristics provided by LPS for their first home.

As recommended by Equifax, we only consider the individuals with an Equifax-provided match confidence score of at least 0.8. To identify the loan that corresponds to the first home purchase (oldest first mortgage) referred to in the consumer credit panel, we identify the oldest first mortgage loan on record by origination year for each matched individual in the CRISM dataset. We eliminate loans that have large discrepancies between origination date and first payment date; loans with payment dates prior to origination dates are dropped as well as loans with first payment dates over two years after the origination date. We then compare the year of first home purchase in the consumer credit panel with the origination year of the oldest loan on record for the matched individual. We drop all loans with over a two year difference between the year of first home purchase and the origination year of the oldest loan record in CRISM. Because the CRISM dataset begins in 2005, we can only observe loans that are still active in 2005 if they originated prior to 2005.

As a result, not all individuals with a match in the CRISM dataset will be linked to a mortgage account in the LPS data. Of the 15,872 homebuyers in the original sample from the consumer credit panel, we are able to find mortgage characteristics for 7,437 individuals about 47% of the homebuyers.

4. House Prices and the Timing of First Home Purchase

In this section, we empirically examine how house price growth affects the housing demand over the life cycle using micro-level data.

4.1 HOMEOWNERSHIP RATE

The Equifax data do 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 the individual's oldest mortgage account. Based on this information, we derive a measure for whether or not individuals 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 homeowner in a given year t if that person purchased a home in year t *or earlier*. Accordingly, we define the homeownership rate in year t as the fraction of individuals in year t who have ever purchased a home by year t. Specifically, let B_t 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 follows:

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

We calculate the homeownership rate by age in a similar fashion. Specifically, let B_{at} 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 rate for age group a in each year, H_{at} , can be calculated as follows:

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

Figure 1(a) plots the aggregate homeownership rate H_t over the time period 1999-2012. Over the sample period between 1999 and 2012, the fraction of homeowners varies between 44% and 47%, with an average of around 46%. The magnitude is similar to Mian and Sufi (2011) — they use the same Equifax data but with a slightly different definition for homeownership rate and find that the homeownership rate to be around 40%.[8]

Figures 1(b)-(d) plot the home ownership rates for three age groups (age 25-34, 35-44, and 45-60) over the sample period[9]. Two things are worth noting. First, the fraction of individuals ever owning a home for the young group (25-34) is lower in level than that for the two older age groups (35-44 and 45-60). Second,

[8] Both magnitudes are, however, lower than the homeownership rate measured using other data (such as the U.S. Census, *Current Population Survey* (CPS) or *American Housing Survey* (AHS)). For example, Fisher and Gervais (2011) use the Census data to examine homeownership trends and find that over 1996-2007, the homeownership rate for households headed by owners of age 25+ rose from roughly 65% to 68%. There are two main reasons for the discrepancy. First, our homeownership rate is based on individual-level data, while the homeownership constructed by Census data is at the household level. Second, Fisher and Gervais (2011) count homeownership for households with owner heads aged 25 and older, while we consider all individuals between age 18 and 60. By including people younger than 25 (who do not usually own homes) and excluding the elderly (who often do), our sample would have a lower fraction of homeowners.

[9] The group for age 18-24 is not included because its average homeownership (round 5%) is generally much lower than that of the other groups

homeownership varies over time and the variation is greater for the young group than for the older groups. For example, after steadily increasing in the early 2000s across all ages, the ownership rate trended down for the young group in the second half of the decade while it remained relatively stable among the two older groups. Over this period, housing prices also varied greatly. In the next section, we examine whether and how (expected) house prices affect people's decisions about when to enter the housing market.

[Insert Figure 1]

4.2 HAZARD RATE OF FIRST HOME PURCHASE BY AGE AND HOUSING PRICE GROWTH

For our paper, the key concept to examine is the hazard rate of the 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 she never bought a home in a previous year? Using cross sections from the data, we can compute this conditional probability (the hazard rate) of the 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 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 initially compute the hazard rate of first home purchase by age for each city-year cell as described before. Then we group the city-year cells into four 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 2 presents the average hazard rate for each of the four groups, ranging from the lowest- to the highest-quartile of housing price growth.

Figure 2 shows that the hazard rate of home purchase over the life cycle 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 afterward. This suggests that people tend to purchase their first home at young 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 rate to compute and compare the counterfactual cumulative distribution functions (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 through 60, we calculate the probability that an individual would purchase a home at or before that age as:

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

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 four HPI growth groups. Figure 3 presents the resulting (counterfactual) CDFs.

Figure 3 shows that the distribution of age at first home purchase for the high HPI growth city-year groups lies uniformly above and to the left of the distribution for the low HPI growth city-year groups. In other words, the former distribution stochastically dominates the latter. Figures 2 and 3 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 consistent with the notion that the accelerating effect of rising house prices through the house price expectation channel more than offsets their decelerating effect through the liquidity constraints channel.

Note that so far our comparisons are made between city-year groups with high versus low housing price growth and use 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 individual's decision to purchase a home), we turn to multivariate analyses using individual-level panel data in the next section.

4.3 ESTIMATION OF HAZARD RATE – MULTIVARIATE ANALYSIS

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

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

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, the individual-level time-varying variable X'_{iact} (risk-score), and variables proxying 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 individual's first home purchase is then

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

where $\Phi(\cdot)$ is the CDF of a standard normal distribution.

Table II reports the estimates from the full sample of CBSAs over the period 1999-2012. Our main result can be found in Column 1. The coefficient on HPI growth is positive and significant. If we assume individuals form their expectations of future house prices based on recent housing price growth, the result suggests that all else being equal, at a given age, individuals who live in cities with expected high future house price appreciations are more likely to buy a first home than their counterparts in areas with expected low future housing prices.

[Insert Table II]

There is distinct life-cycle pattern in the hazard rate 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 (in Figure 2 — the hazard rises sharply from the late 20s, peaks around 30, and then declines gradually afterward). 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, since 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 expectations of future prices. For example, cities that experienced rapid house price appreciation 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 more and earlier entrants 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 individual's 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 rate of home purchase indeed increases monotonically with HPI growth.

Since 3-year HPI growth might also reflect growth between $t-1$ and t, this period may be when the first home purchases occur. One concern is that increases in home purchases may cause house prices to go up during year t, leading to reverse causality. To alleviate 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 results (Column (5)) are largely unchanged. It suggests that our result is not driven by the effect of increasing house demand on current house prices.

4.4 EVALUATE THE MAGNITUDE OF THE HPI EFFECT

While the estimated coefficients from the model show that HPI growth has a positive effect on the hazard rate of an individual's 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 the following: 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: (1) we assume that individuals in our sample have always lived in the cities with the lowest HPI growth (the bottom quartile) versus (2) we assume that individuals in our sample have always lived in cities with highest HPI growth (the top quartile).

[Insert Figure 4]

Under each scenario, we use the estimates from the probit model in Column 4 of Table II to predict for each individual the hazard rate of home purchase at each age from 18 through 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 rates across all individuals. Figure 4, Panel (a) presents the average hazard for the top quartile and bottom quartile separately. Panel (b) presents the proportional margional difference. The figure shows a V-shaped heterogeneous effect of house price growth — when house price growth increases, the relative increase in the hazard rate of buying first home is larger for young individuals than for the middle-age group. This result is consistent with our model's prediction that individuals have a higher probability to purchase home at a young age than at middle age.

Similarly, under each scenario, we also use the predicted hazard rates to estimate the CDF for each individual at each age and then take the average across all individuals. Figure 4, Panel (c) presents the average CDF from the experiment.

According to the counterfactual distributions in Panel (c) of Figure 4, the difference in the age at which the median individual becomes a first-time home buyer (that is, the age by which half of the population has bought a home) is about 5 years: First-time home buyers are 5 years younger under scenario 2 than under scenario 1 (39 versus 44 years).

Note that as we only have estimates of the hazard and survivor functions for individuals up to 60 years old, 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 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 versus 32 years). Note the difference in the conditional median is smaller than the difference in the unconditional median we estimated earlier.

4.5 ROBUSTNESS TESTS

One might still be concerned about the potential endogeneity of housing price growth if the controls in the model do not capture all the unobservables that affect the local housing demand through changes in fundamentals (such as demographics or economic conditions). To partially address this issue, we instrument local house prices with national house prices interacted with local house supply elasticity (Saiz, 2010). The rationale for the instrument is based on the intuition that when there is an aggregate shock to housing demand (say, lower interest rates), house prices could rise by different degrees across areas depending on the supply response. For a given increase in demand, prices might rise more in areas where the supply is less elastic. Specifically, we create an instrument for our 3-year change in HPI growth variable by using a national measure of house price growth (also taken from CoreLogic) multiplied by the elasticity of the housing supply in each CBSA.[10] We then estimate an ivprobit model for the hazard rate using a subsample of CBSAs for which the supply elasticity information is available. The results for the first and second stage are listed in Panel A of Table III, which continue to show a positive effect of housing price growth on the likelihood of home purchase. In fact, the coefficient is larger in magnitude than the probit baseline (Column 1 of Panel A), although it is also less precisely estimated.

Furthermore, as our data on individual credit scores, addresses, etc., only goes back to 1999, we have to limit the estimation 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 by only keeping young individuals (18-25 years old when they enter our data). Since this cut the sample size dramatically, we used a 5% primary random sample, rather than the 1% primary sample we use for the rest of our analysis. The results remain the same, as shown in Panel B of Table III.

Another concern is the confounding effect of migration. Geographic patterns of industry agglomeration (e.g., a Silicon Valley effect) may lead to spatial groupings of young people receiving unusually large incomes, thereby fueling both rapid house price increases and sooner-than-average first home purchases. To address this concern, we also restrict our sample to individuals who have been in the same CBSA for at least three years prior to the purchase of the house. The results remain the same, as reported in Panel C of Table III.

[Insert Table III]

4.6 HOUSE PRICE GROWTH AND EXTRAPOLATIVE EXPECTATIONS

The results show that individuals purchase their first homes at young ages when they have lived in places experiencing high house price growth in the past few years. The direction of the empirical relationship between housing price and age at purchase runs contrary to the prediction from the liquidity constraint hypothesis, but is consistent with the expectation channel hypothesis — rising past price growth leads to expectations of higher future price growth and thus accelerates home purchases at early ages. Case et al.

[10] Saiz (2010) provides estimated elasticities for 95 of the largest CBSAs, so we estimate the model using data from these CBSAs (which cover a little less than half of the observations in our survey data).

(2012) conducted an annual survey of US home-buyers in four metropolitan areas from 2003 through 2012. The survey elicits quantitative estimates of expected future house price growth, specifically, homebuyers' expectations of the change of the value of their homes in the next year and ten years. They show that the expected change in home prices is positively correlated with actual lagged price changes based on the S&P/Case-Shiller Home Price Index. Specifically, they run a simple regression of the reported expected one-year change in house prices on one-year lagged price changes and find a statistically significant coefficient of 0.23 and $R^2 = 0.73$. Similarly, when we regress the reported expected one-year price change from their survey on the actual average annual price change in the past three years based on the CoreLogic home price index as used in our paper, we find a slope coefficient of 0.29 (with p-value$<$0.01) and R^2 of 0.70, as shown in Table IV. This provides additional supportive evidence that individuals form expectations of high future price growth largely based on past price growth, and corroborates the expectation channel hypothesis that links past price growth with individuals' timing of homeownership.

[Insert Table IV]

5. Mortgage-supply Explanation vs Housing-demand Explanation

It is, however, worth noting that the extrapolative expectation may apply not only to home buyers, but also to credit suppliers (as discussed at the end of the conceptual model section). Given expectations of high future price growth, credit suppliers may lower lending standards and thus relax buyers' borrowing constraints. The relaxation of borrowing constraints may be confounded with the accelerating effect of the extrapolative expectation from the home buyers, because it could also lead to a larger increase in the hazard rate of first home purchases for the young group than relative to the middle-age group, if the young group is more financially constrained than the middle-age group, other things being equal. In this section, we conduct some tests to separate these two competing effects. First, we show that the accelerating effect of housing price growth on first home purchases still exists during a subperiod or for a subsample of individuals for which there was less variation from the mortgage supply side. Next, we provide evidence related to house size and down payment and show that the housing-demand story is a necessary and complementary explanation to explain the evidence all together.

5.1 SUBPERIOD FROM 1991 TO 1999

To separate the housing demand versus mortgage supply explanations, we extend our sample back to the 1990s, the period when the mortgage supply side forces vary less across different regions. If the mortgage-supply side is the main force driving our results and the housing-demand plays very little role, we should expect no accelerating effect of house price growth on first-home purchases in this period since mortgage supply was unlikely to be dramatically different across regions.

Table V reports the results from reestimating our main specifications for the 90s sample. We are able to extend the panel back to the 90s because we identify individual's first home purchases by the person's oldest mortgage account. The sample is restricted to only observations from 1991 to 1999. Individuals exit the sample after the year of their first home purchase. We include the same set of control variables as in our main analysis, except the time-varying credit score, because the credit score from the 90s is not available in the Equifax dataset.

The results remain the same—despite the mortgage supply remaining stable in the 90s, we still find that the hazard rate of first-home purchase increases monotonically with HPI growth. The evidence corroborates the hypothesis that home buyers develop extrapolative expectations on future house price growth, and purchase house earlier in their life cycle when they witness experience high house price growth in the past few years.

[Insert Table V]

5.2 NON-SUBPRIME INDIVIDUALS

To further test the two explanations, we also repeat our analysis on the non-subprime borrowers. It is well documented that credit supply increases were mainly concentrated in the subprime areas (Mian and Sufi, 2009). Dell'Ariccia et al. (2012) also document that the deterioration of lending standards was more of a subprime market phenomenon, and much weaker or absent in the prime mortgage market.

Therefore, to isolate the housing-demand explanation for first home purchases at young ages in areas with high house price growth, we focus our analysis on a subsample of non-subprime borrowers who were less likely to be much affected by the credit supply change. Specifically, we look at a subsample of borrowers with FICO scores above 660 (Keys et al., 2010) when they enter our sample.

The results, as shown in Table VI, are largely unchanged and deliver a similar pattern of V-shaped heterogeneous effect of house price growth as previously seen in Panel (b) of Figure 4. The results provide support for the housing-demand explanation, because the mortgage-supply explanation is less applicable for the non-subprime individuals given the absence of the deterioration of lending standards.

[Insert Table VI]

5.3 EVIDENCE ON HOUSE SIZE AND DOWN PAYMENT

One might argue that the laxing of lending standards to include more risky borrowers is only part of the story for the expansion of the mortgage supply. Given high expectations of future house price growth, lenders may also be willing to ask for lower down payment (higher LTVs) which would relax borrowing constraints for individuals, especially for young buyers. In that case, we would expect to see a positive effect of house price growth on LTV, especially at the upper tail of the LTV distribution. To test this, we merge the Equifax and the LPS datasets and run quantile regressions of LTV on past house price growth (both level and quartile dummies).

The results are presented in Table VII. Column 1-2 and 3-4 are estimated separately for each conditional quantile: 0.5 and 0.75. As the table shows, the coefficients are either insignificant or significantly negative, indicating that the LTV does not increase with the past house price growth across the distribution, including the upper tail, of LTV. The evidence is not consistent with the mortgage-supply side explanations, lenders lower down payment given expectations of high future house price growth, since it would suggest a higher rather than lower LTV ratio associated with higher past price growth.

[Insert Table VII]

Despite the evidence of unchanged or even lower LTV in areas with high house price growth, one might still think that the required down payment is lowered (looser borrowing constraints) in those areas. In this subsection, we provide further evidence related to house size and down payment and show that the morgage-supply explanation, though might be applicable, can not be the only story to explain the accelerating effect of housing price growth on the first-home purchases.

5.3.1 Optimal House Size

If young buyers are financially constrained, when they form higher expectations about future house prices and increase their demand for housing, they are likely to purchase a smaller home and enter the housing market earlier. In contrast, if the mortgage supply-side explanation is the only story, young buyers borrowing constraints are relaxed, and we would not expect them to reduce house sizes.

To test this hypothesis, we regress house size (proxied by the price of the purchased house deflated by the house price growth) on house price growth, and then compare the response of house size to house price growth between young buyers (age$\leq$30) and old buyers (age$>$30). If the mortgage supply-side explanation is the only channel, we would expect no reduction in the size of houses bought by either young or old buyers. If anything, we would expect house sizes become larger in areas with higher house price growth, especially for young buyers.

Table VIII present the results of regressing house size on quartile dummies of house price growth for young and old buyers separately. The two age groups exhibit very different patterns. Young buyers tend to choose smaller houses in areas with higher house price growth, whereas the house size chosen by older buyers shows an increasing pattern with house price growth. The downsizing pattern of young buyers who experienced high house price growth is more consistent with the demand-side explanation, and certainly cannot be explained only by the mortgage supply-side story. The evidence also provides additional insight that the accelerating effect of house price growth on first home purchases may have differential impact on the market for houses of different sizes, with more amplification for smaller houses. For example, Agarwal et al. (2014) show that the condominium loan market experienced a 15-fold increase in origination from 2001 to 2007. The differential impact on the housing market segments may enhance our understanding about the recent financial crisis.

[Insert Table VIII]

5.3.2 Down Payment

Another dimension we examine is down payment. If the mortgage supply-side explanation is the only story to explain the accelerating effect, young buyers (the more financially constrained group) are more likely to purchase a house of the same size only if they could lower the total amount of down payment.

We regress down payment on past house price growth (controlling for house size) for the young and old buyers separately. If the demand channel does not exist and the accelerating effect could be fully explained by the mortgage supply-side story, we would expect that, holding house size constant, down payment will decrease, especially for young buyers, as house price growth increases. However, as shown by Table IX, both groups pay higher (rather than lower) down payment in areas with higher past house price growth. The pattern of increasing down payment for a given house size, which likely come from the higher house prices, rules out the possibility that the supply-side story alone could explain the accelerating effect documented in our paper.

[Insert Table IX]

5.4 SUMMARY

To sum up, while we cannot completely rule out the mortgage supply-side explanation, there is strong evidence that the housing demand explanation plays a crucial role in explaining the accelerating effect of house price growth on first-home purchases. First, we find that the accelerating effect still appear during the 90s when the mortgage supply varies little across different regions or for the non-subprime individuals who were less likely to be much affected by the credit supply change. Furthermore, we find that the mortgage-supply explanation alone can not explain the results related to house size and down payment. The house-demand story is a necessary mechanism for one to explain all evidence documented in the paper. We view the house-demand side story as complementary, rather than alternative, to the mortgage-supply side explanations.

6. Conclusion

In this paper we investigate how house price growth affects individuals' timing of the first home purchase over the life cycle. A priori, there might be two offsetting channels that lead to two opposite hypotheses.

The first one is the *liquidity constraint channel hypothesis*. Larger past price increase may lead to a higher level of current house price. Higher current prices may postpone home purchases if individuals are financially constrained and thus less likely to be able to meet the larger down payment requirement. The second one is the *expectation channel hypothesis*. Higher recent price growth may also lead households to expect house prices to rise faster in the near future. Expectations of higher future price may prompt individuals to prepone first home purchases because of investment or hedging incentives.

We use a unique panel data set covering the period 1999-2012 that allows us to track the same individual over these 13 years and observe her decision to buy a house. We show that when house prices are rising, individuals tend to buy houses earlier in their life cycle, which is more consistent with the expectation hypothesis, i.e., all else being equal, expectations of higher future house prices accelerate house purchase in early ages. By exploiting the time-varying cross-sectional variations in house price growth across metropolitan areas while controlling for other potentially confounding demand and supply factors, we find that individuals accelerate their probability of buying a house in metropolitan areas with the highest quartile house price growth relative to metropolitan areas with the lowest quartile house price growth. Specifically, we find that the age at which the median individual becomes a first-time home buyer in the population goes down (buying at an earlier age) by 5 years when we compare individuals who live in the highest house price growth quartile with those who live in the lowest house price growth quartile. The accelerating effects are likely to be driven by extrapolative expectations of both mortgage suppliers and home buyers. We provide evidence that the accelerating effect exists even in cases when the mortgage-supply story is likely to be absent, and the mortgage-supply story alone can not explain the evidence related to house size and down payment. Therefore, the housing demand explanation plays a crucial role in explaining the accelerating effect of house price growth on first-home purchases.

Our key contributions to the existing literature can be summarized as follows. First, we are the first to use the FRBNY CCP data set to study housing demand over the life cycle. This data set is a large panel, with little measurement error, following individuals over a fairly long time horizon with detailed geographical information. We also use a survival type of analysis that explicitly accounts for censoring. This also allows us to study the entire age distribution of potential home buyers (as some will never buy a home). Second, the prior housing literature focuses mostly on the level of homeownership and not the timing of homeownership. We show that the timing is impacted by local past house price growth — individuals prepone the first house purchase when they live in places that have experienced high local house price growth in recent years. Besides providing implications for the aggregate effect on the housing market, the shift in timing may also provide additional insight about the differential impacts across various submarkets. For example, our evidence that, when house price grows fast in recent years, individuals purchase their first houses earlier with smaller sizes, suggests that the recent house price growth has a larger amplification effect on the market for smaller houses.

A Appendix

A.1 CONCEPTUAL MODEL

A.1.1 Case 1: Stable house price ($\lambda = 0$, i.e., $P_0 = P_1 = P_2 = P$)

In this case, individuals believe that house prices will stay at the same level for all periods. There is no investment incentive to purchase a house. The benefit of buying a house comes only from the extra utility of owning a house at time 1. Since buying a house requires a down payment and thus reduces 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, individuals would choose the optimal consumption level as $C_0^{NC} = C_1^{NC} = \frac{Y}{3}$. But since individuals have to make a down payment for a house purchase, they may not achieve the unconstrained optimum. Specifically,

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

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

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

4. If $\frac{Y}{2} - dP < 0$ and $Y - dP \geq 0$, individuals can not afford the down payment at time 0 any more, but by 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 $Y - dP \geq 0$, lifetime income is too low to afford the down payment, so the individual 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 a house in their middle age.

[Insert Figure A.1]

A.1.2 Case 2: Downward house price ($\lambda < 0$)

In this case, at time 0, individuals 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 that individuals 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 at 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 at 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 is 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.1.3 Case 3: Upward house price ($\lambda>0$)

In this case, at time 0, individuals expect that house prices will increase in the next period. They may consider buying homes at a young age because (1) buying a home at time 0 could lead to investment opportunities (investment motives for homeownership) and (2) some may not be able to afford the high house price at time 1 if not buying at time 0 (consumption motives of homeownership). We will discuss individuals' decisions in different regions of $\frac{\lambda P}{Y}$ and $\frac{dP}{Y}$ in more detail. Then, to separate out 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), individuals cannot afford a house 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), individuals cannot afford at time 0 but can afford a house 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), individuals can afford a house at time 0, but will not be able to afford one if they wait until time 1 to buy one. 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), individuals can afford a house in both periods. We proceed as follows: (1) solve the optimal consumption plans given the individual chooses to buy a home at time 0 and at time 1; and (2) compare the utilities based on the two conditional optimized consumption plans and then decide which period is the best time to buy a home.

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

   (a) If $\frac{Y}{2} - dP \geq \frac{1}{3}(Y + \lambda P)$ (Area D1 in Figure A.2), individuals can maintain the unconstrained optimal consumption and at the same time buy a house at time 0. So their best choices are to buy a house 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), individuals are not able to maintain $C^{NC}_{0|\tau=0}$ if buying a house at time 0. So the corner solution gives that, at time 0, individuals consume $C^*_{0|\tau=0} = \frac{Y}{2} - dP$. To further split the region, we do the following:

      - 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 individuals cannot borrow from the future other than buying a 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), individuals 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 individuals buy a home at time 1, the unconstrained optimal consumption plan is as follows: $C^{NC}_{0|\tau=1} = C^{NC}_{1|\tau=1} = \frac{Y}{3}$.

(a) If $\frac{Y}{3} \geq d(1+\lambda)P$ (Area D2 and D4 in Figure A.2), even though individuals buy a house 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), individuals cannot achieve the unconstrained consumption level if they buy a house 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 when buying a home at time 0 and buying a home 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 individuals choose to buy a house 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, individuals buy a home at their young age (time 0). They are willing to sacrifice their consumption at a young age to either profit from the investment opportunity or hedge to assure they could become home owners during their middle age.

A.1.4 Testable Predictions

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

Proposition 2. *All else being equal, the probability of buying a home at time 0,* $\Pr(\tau^* = 0)$*, decreases with the current price P, but increases with price expectation parameter λ.*

Proof. When $\lambda \leq 0$, individuals always prefer to buy a house at time 1 or are indifferent about the timing of the purchase (time 0 or time 1), so $Pr(\tau^* = 0) = 0$;

When $\lambda > 0$, the probability of buying a home 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 $\lambda \geq 1$. 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 $\frac{dP}{Y^c} = K^c$), individuals choose to buy homes at time 0. When $0 < K \leq \frac{1}{6}$ (i.e. $Y \geq 6dP$), $L(K) = \frac{\lambda}{d}K \geq 0$, we have already known that individuals will also choose to buy homes at time 0.

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

Therefore, $\frac{\partial \Pr(\tau^*=0)}{\partial P} < 0$, $\frac{\partial \Pr(\tau^*=0)}{\partial \lambda} > 0$.

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 $K \in (0, \frac{1}{2}]$, individuals choose to buy a home at time 0. So, $\Pr(\tau^* = 0) = \int_{2\mathrm{dP}}^{\infty} dF$. Therefore, $\frac{\partial \Pr(\tau^*=0)}{\partial P} < 0$, $\frac{\partial \Pr(\tau^*=0)}{\partial \lambda} = 0$.

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

One caveat is that our model only illustrates the decision-making of individuals and abstracts away the action of mortgage credit suppliers. Specifically, we assume a fixed minimum down payment ratio, which may vary with market conditions in a general equilibrium setting. If instead we assume mortgage credit suppliers would relax borrowing constraints (i.e., lower the down payment ratio d) given expectations of high future house prices, there will be two scenarios:

(1) If home buyers have negative or zero price growth expectation ($\lambda \leq 0$), they will prefer to buy at time 1 or be indifferent about when they buy (at time 0 and at time 1). In this case, relaxing borrowing constraints does not influence the timing of individual's first-home purchases.

(2) If home buyers share similar positive expectations of house price growth ($\lambda > 0$) as the credit suppliers, there will be discuss two cases as shown above: When $0 < \lambda < 1$, $Pr(\tau^* = 0) = \int_{Y^c}^{\infty} dF$ and $\frac{\partial Y^c}{\partial \lambda} = -\frac{\partial F(K^c)/\partial \lambda}{\partial F(K^c)/\partial Y^c} = -\frac{K^c/d}{F'(K^c)\cdot(-\frac{dP}{{Y^c}^2})} < 0$. It's easy to get that $\frac{\partial Y^c}{\partial \lambda}$ decreases when d decreases. Therefore, $Pr(\tau^* = 0)$ increases when d decreases. When $\lambda \geq 1$, $\Pr(\tau^* = 0) = \int_{2\mathrm{dP}}^{\infty} dF$, so $Pr(\tau^* = 0)$ also increases with d decreases. To sum up, when $\lambda > 0$, relaxing borrowing constraints will then amplify the acceleration effect of home buyers' price growth expectation on first home purchases.

References

Agarwal, S., Y. Deng, C. Luo, and W. Qian (2014). The hidden peril: The role of the condo loan market in the recent financial crisis. *Working Paper*.

Attanasio, O. P., R. Bottazzi, H. W. Low, L. Nesheim, and M. Wakefield (2012). Modelling the demand for housing over the life cycle. *Review of Economic Dynamics 15*(1), 1–18.

Carroll, C. D. (1997). Buffer-stock saving and the life-cycle/permanent income hypothesis. *Quarterly Journal of Economics* (112), 1–55.

Case, K. E., R. J. Shiller, and A. K. Thompson (2012). What have they been thinking? homebuyer behavior in hot and cold markets. *Brookings Papers on Economic Activity*, 265–298.

Cocco, J. F. (2004). Portfolio choice in the presence of housing. *Review of Financial Studies 18*(2), 535–567.

Deaton, A. (1991). Savins and liquidity constraints. *Econometrica* (59), 1221–1248.

Dell'Ariccia, G., D. Igan, and L. Laeven (2012). Credit booms and lending standards: Evidence from the subprime mortgage market. *Journal of Money, Credit and Banking 44*(2-3), 367–384.

Fetter, D. K. (2013). How do mortgage subsidies affect home ownership? evidence from the mid-century gi bills. *American Economic Journal: Economic Policy 5*(2), 111–147.

Fisher, J. D. and M. Gervais (2011). Why has home ownership fallen among the young? *International Economic Review 52*(3), 883–912.

Han, L. (2010). The effects of price risk on housing demand: Empirical evidence from u.s. markets. *Review of Financial Studies 23*(11), 3889–3928.

Keys, B. J., T. K. Mukherjee, A. Seru, and V. Vig (2010). Did securitization lead to lax screening? evidence from subprime loans. *Quarterly Journal of Economics 125*(1), 307–362.

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

Lee, D. and W. van der Klaauw (2010). An introduction to the frbny consumer credit panel. *FRBNY staff Reports*.

Mian, A. and A. Sufi (2009). The consequences of mortgage credit expansion: Evidence from the u.s. mortgage default crisis. *Quarterly Journal of Economics 124*(4), 1449–1496.

Mian, A. and A. Sufi (2011). House prices, home equity-based borrowing, and the us household leverage crisis. *American Economic Review 101*(5), 2132–2156.

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

Piazzesi, M. and M. Schneider (2009). Momentum traders in the housing market: Survey evidence and a search model. *American Economic Review Papers and Proceedings*, 406–411.

Qian, W. (2013). Why do sellers hold out in the down market? an option-based explanation. *Real Estate Economics 41*(2), 384–417.

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

Sinai, T. and N. S. Souleles (2005). Owner-occupied housing as a hedge against rent risk. *Quarterly Journal of Economics 120*(2), 763–789.

Stroebel, J. and J. Vavra (2015). House prices, local demand, and retail prices. *Working Paper*.

Figure 1: Probability of Homeownership (Full sample and by age group), 1999-2012
The figure presents the homeownership of the full sample and by three age groups (25-34, 35-44, 45-60) for each year from 1999 through 2012. The homeownership rate in year t is defined as the fraction of individuals in year t who have ever purchased a home by year t. The time individuals bought their first home is identified by the age of their oldest mortgage account.

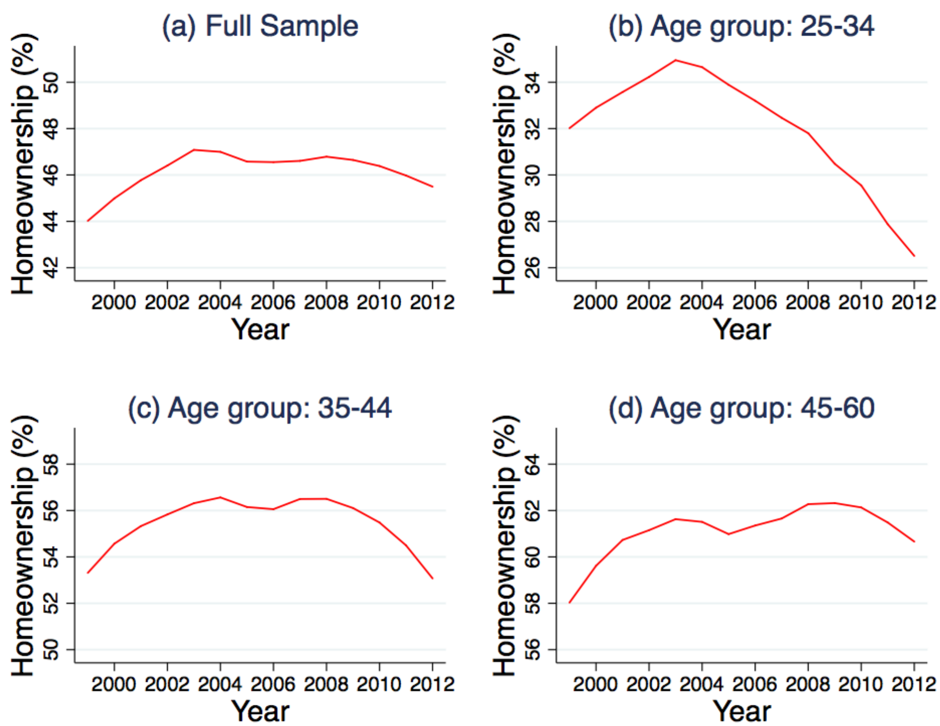


Figure 2: Hazard Rate of Buying First Home by Age

The figure presents the average hazard rate of buying a first home for four city-year groups sorted by lagged three-year CoreLogic Home Price Index (HPI) growth. For each city-year cell, the hazard rate is computed as the total number of individuals of a given age within that city-year cell who purchase their first home divided by the total number of individuals of that given age within that city-year cell who have never bought a home before. The hazard rate is then averaged across city-year cells within each group sorted by lagged three-year HPI growth, weighted by the number of people in the city-year cell.

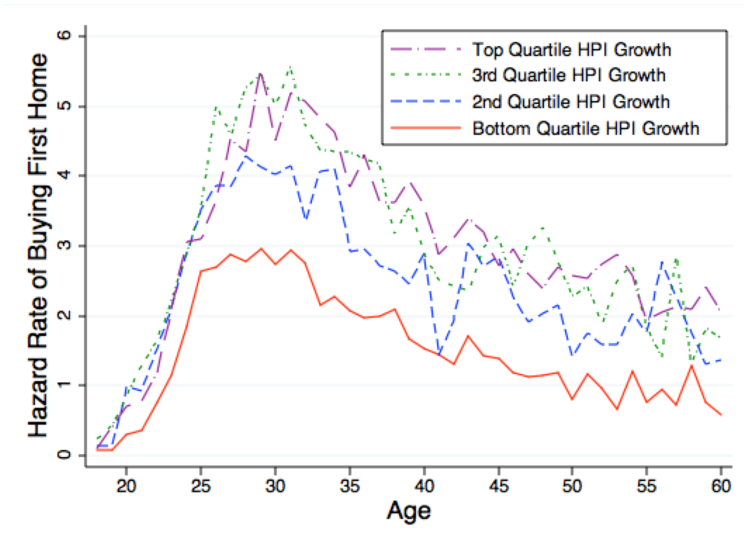


Figure 3: Cumulative Distribution of Homeownership by Age
The figure presents the cumulative probability of purchasing a home at or before a given age for four groups sorted by the lagged three-year HPI growth. The cumulative probability is first constructed by city-year cell and then averaged across all city-year cells within each group, weighted by the number of people in the city-year cell.

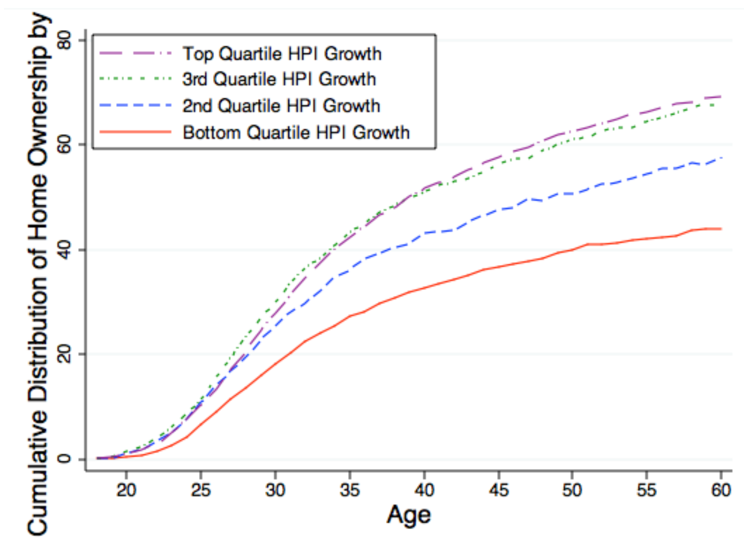
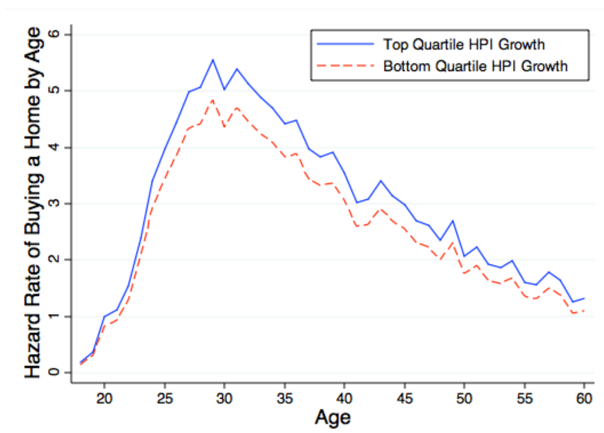
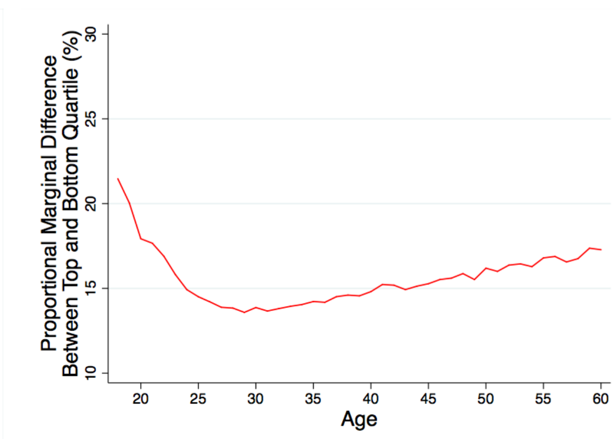


Figure 4: The HPI Effect by Age (Estimated by Probit Regression)
The figure presents the hazard rate of first home purchase at each age between 18 and 60 predicted by the estimates from the probit model in Column (4) of Table II. For each prediction, the individuals are assumed to have always lived in the cities with either the highest HPI growth (the top quartile) or the lowest HPI growth (the bottom quartile), and only the HPI growth and age variables are allowed to vary at their hypothetical values while all other variables are kept at their actual values in the data. The predicted hazard is then averaged across all individuals within each group. Panel (a) presents the average predicted hazard rate for the top quartile and bottom quartile separately. Panel (b) presents the difference of the average predicted hazard rates between these two groups normalized by the sum of the hazard rates of these two groups. Panel (c) presents the predicted cumulative probability.

(a) Hazard Rate of Buying First Home by Age



(b) Proportional Marginal Difference



(c) Cumulative Distribution of Homeownership by Age

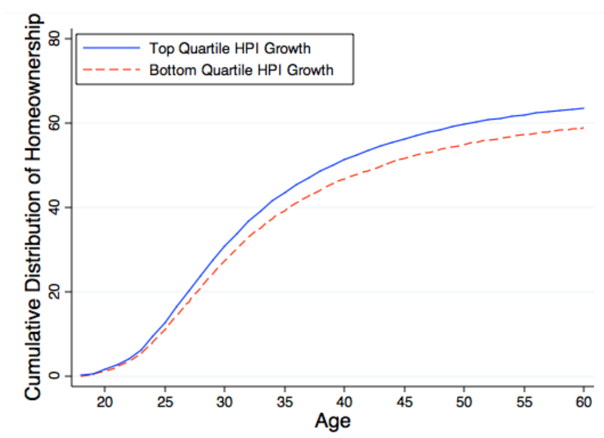


Table I: Summary Statistics

Panel A reports the summary statistics for the 1% sample of the primary-individual data from the FRBNY (Equifax) Consumer Credit Panel Dataset. Our sample covers the period from 1999 through 2012 and only include individuals with age between 18 and 60 years old. Panel B reports the summary statistics for the house price growth constructed by CoreLogic Home Price Index (HPI). Panel C reports the summary statistics for CBSA-level variables constructed by county-level data from the *Quarterly Census of Employment and Wages* (QCEW) compiled by the BLS.

	Obs	Mean	Std	25th	Median	75th
A. FRBNY (Equifax) Consumer Credit Panel						
Pooled individual-year obs:						
Age	1184591	39.91	11.55	30.00	40.00	50.00
Credit Score	1096090	666.87	106.05	589.00	677.00	758.00
Individual obs:						
Age at first-home purchase	16689	35.36	10.01	27.00	33.00	42.00
Credit score range for individuals over time (High-Low)	116128	124.91	86.83	57.00	115.00	182.00
Consumer credit:						
Credit card balance	1182788	3763.15	9254.63	0.00	639.50	3557.75
Auto loan balance	1182788	4643.74	9855.08	0.00	0.00	6060.00
Delinquency Rate for accounts with nonzero balances (1 = 90+ days past due and severe derogatory or bankruptcy; 0 = otherwise):						
Credit card	793849	0.19	0.39			
Auto loan	427830	0.09	0.29			
Mortgage	104176	0.08	0.27			
B. CoreLogic House Price Index						
Pooled CBSA-year obs:						
3-year average HPI growth (%)	13351	2.90	6.05	-0.79	3.61	6.11
CBSA obs:						
HPI growth range for CBSA over time (High-Low) (%)	960	20.64	11.76	13.36	16.34	24.88
C. BLS: Employment and Wage Data						
Annual average employment growth (%)	960	0.31	1.28	-0.40	0.28	0.94
Annual average weekly wage growth (%)	960	2.90	0.77	2.49	2.83	3.21
Annual average quarterly growth in number of establishments (%)	960	0.69	1.09	-0.04	0.55	1.27
Annual unemployment rate (%)	960	6.38	1.87	5.22	6.16	7.31

Table II: Hazard Rate of Buying the First Home (Probit Regression)

The table estimates the hazard rate of buying the first home using a probit binary choice model for first home purchase. The sample is based on the 1% sample of the primary-individual data from the FRBNY (Equifax) Consumer Credit Panel Dataset. Individuals exit the sample after they year of their first home purchases. The main variable is the house price growth (HPI growth) and the quartile dummies sorted on the HPI growth among the pooled CBSA-year observations. Control variables are the time-varying individual-level credit score, CBSA-level employment and wage variables, year, age, and CBSA effects. Standard errors are clustered at the CBSA level and are shown below the coefficient estimates. *10%, **5%, ***1% significance.

	(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	0.342***	0.303***	0.202*		
	(0.108)	(0.110)	(0.116)		
2nd quartile HPI growth (CBSAs)				0.0576***	0.0423**
				(0.0180)	(0.0177)
3rd quartile HPI growth (CBSAs)				0.0505**	0.0574***
				(0.0210)	(0.0206)
4th quartile HPI growth (CBSAs)				0.0697***	0.0731***
				(0.0216)	(0.0202)
Credit score (time-varying)	0.00274***	0.00274***	0.00274***	0.00274***	0.00274***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Annual average employment growth		0.327	0.184	0.175	0.210
		(0.239)	(0.257)	(0.256)	(0.258)
Average weekly wage growth		0.411	0.390	0.384	0.378
		(0.253)	(0.265)	(0.264)	(0.266)
Average quarterly growth in number of establishments		-0.126	-0.149	-0.115	-0.137
		(0.196)	(0.194)	(0.193)	(0.194)
Unemployment rate			-1.165**	-1.188**	-1.226**
			(0.584)	(0.586)	(0.565)
Year dummies	Yes	Yes	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes	Yes	Yes
CBSA dummies	Yes	Yes	Yes	Yes	Yes
Observations	524,168	524,168	523,768	523,768	523,768

Table III: Robustness Tests for Hazard Rate of Buying the First Home
In Panel A, we instrument the 3-year local house price growth with national house price growth interacted with local house supply elasticity (Saiz, 2010). The sample only includes the CBSAs which could be matched with supply elasticities reported in Saiz (2010). For comparison, we feature Column (1) reporting the hazard rate estimated by our baseline probit model for the same subsample. Panel B estimates the hazard rate of buying the first home using a probit binary choice model for first home purchase. The sample is based on the 5% sample of the primary-individual data from the FRBNY (Equifax) Consumer Credit Panel Data set, and only includes individuals who enter the sample between 18 and 25 years old. Individuals exit the sample after the year of their first home purchase. The main variable is the quartile dummies sorted on the HPI growth among the pooled CBSA-year observations. Control variables are the time-varying individual-level credit score, CBSA-level employment and wage variables, year, age, and CBSA effects. Panel C restricts the sample to individuals who lived in the same CBSA for at least three years. Standard errors are all clustered at the CBSA level and are shown below the coefficient estimates. *10%, **5%, ***1% significance.

Panel A: Instrumental Variable Estimation

	(1)	(2)	(3)
	Probit (Subsample)	First Stage	IV Probit
Left-hand side variable:	Hazard Rate	3-year HPI growth from $t-3$ to t	Hazard Rate
HPI growth measure:	3-year HPI growth		Instrumented 3-year HPI growth
HPI growth	0.233	-0.278***	0.525*
(Instrument = Elasticity x National HPI Growth)	(0.162)	(0.0318)	(0.311)
Credit score (time-varying	0.00280***	0.000	0.00280***
	(0.000)	(0.000)	(0.000)
Annual average employment growth	0.502	0.323**	0.408
	(0.443)	(0.142)	(0.468)
Average weekly wage growth	0.389	0.0886	0.354
	(0.415)	(0.0674)	(0.405)
Average quarterly growth in number of establishments	-0.0753	0.142*	-0.103
	(0.266)	(0.0794)	(0.262)
Unemployment rate	-0.413	-2.110***	0.361
	(0.958)	(0.252)	(1.163)
Constant	-4.536***	0.142***	-4.571***
	(0.152)	(0.0111)	(0.156)
Year dummies	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes
CBSA dummies	Yes	Yes	Yes
Wald test of exogeneity p-value		0.263	0.263
Observations	295,746	295,746	295,746

Panel B: Subsample for Individuals Who Enter the Sample Between 18 and 25 Years Old

	(1)	(2)
HPI growth measure:	3-year HPI growth from $t-3$ to t	2-year HPI growth from $t-3$ to $t-1$
2nd quartile HPI growth CBSAs	0.0754***	0.0381*
	(0.0208)	(0.0220)
3rd quartile HPI growth CBSAs	0.0462*	0.0294
	(0.0240)	(0.0260)
4th quartile HPI growth CBSAs	0.0960***	0.0772***
	(0.0259)	(0.0274)
Credit score (time-varying	0.00307***	0.00307***
	(0.000)	(0.000)
Annual average employment growth	0.319	0.387
	(0.249)	(0.251)
Average weekly wage growth	0.548**	0.548**
	(0.255)	(0.255)
Average quarterly growth in number of establishments	-0.0701	-0.0945
	(0.184)	(0.184)
Unemployment rate	-0.565	-0.702
	(0.545)	(0.551)
Constant	-4.959***	-4.920***
	(0.0777)	(0.0800)
Year dummies	Yes	Yes
Age dummies	Yes	Yes
CBSA dummies	Yes	Yes
Observations	650,687	650,687

Panel C: Subsample restricted to individuals who lived in the same CBSA for at least three years

	(1)	(2)
HPI growth measure:	3-year HPI growth from t-3 to t	2-year HPI growth from t-3 to t-1
2nd quartile HPI growth CBSAs	0.0470**	0.0527**
	(0.0219)	(0.0230)
3rd quartile HPI growth CBSAs	0.0377	0.0618**
	(0.0237)	(0.0262)
4th quartile HPI growth CBSAs	0.0511**	0.0804***
	(0.0248)	(0.0251)
Credit score (time-varying)	0.00260***	0.00260***
	(4.17e-05)	(4.17e-05)
Annual average employment growth	0.122	0.146
	(0.267)	(0.266)
Average weekly wage growth	0.432	0.403
	(0.270)	(0.267)
Average quarterly growth in number of establishments	-0.235	-0.240
	(0.204)	(0.204)
Unemployment rate	-1.064*	-0.880
	(0.579)	(0.573)
Year dummies	Yes	Yes
Age dummies	Yes	Yes
CBSA dummies	Yes	Yes
Observations	519,116	519,116

Table IV: Regressions of the Expected One-Year Change in Home Prices on Lagged Actual Price Changes

The table reports the results of regressing the survey-reported expected one-year change in house prices on the 3-year lagged HPI growth. The survey-reported house price growth expectation is from an annual survey of U.S. home buyers in four metropolitan areas from 2003 through 2012, conducted by Case et al. (2012). Standard errors are shown below the coefficient estimates. *10%, **5%, ***1% significance.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: Price change expectations within 12 months					
Survey location:	Alameda County	Middlesex County	Milwaukee County	Orange County	All	All
3-year lagged HPI growth	0.249***	0.353***	0.315***	0.313***	0.287***	0.292***
	(0.0560)	(0.0895)	(0.0477)	(0.0794)	(0.0348)	(0.0338)
Constant	4.651***	2.505***	3.738***	2.820**	3.482***	2.611***
	(0.637)	(0.492)	(0.285)	(1.041)	(0.334)	(0.624)
Observations	10	10	10	10	40	40
R-squared	0.712	0.661	0.845	0.660	0.642	0.695
County FE	No	No	No	No	No	Yes

Table V: Hazard Rate of Buying the First Home (Subsample for 1991-1999)
The table estimates the hazard rate of buying the first home using a probit binary choice model for first home purchase. The sample is based on the 1% sample of the primary-individual data from FRBNY (Equifax) Consumer Credit Panel Dataset. We extended the panel dataset back to 1991 and restricted the sample to only observations from 1991 to 1999. Individuals exit the sample after the year of their first home purchase. The main variable is the quartile dummies sorted on the HPI growth among the pooled CBSA-year observations. Control variables are CBSA-level employment and wage variables, year, age, and CBSA effects. Standard errors are clustered at the CBSA level and are shown below the coefficient estimates. *10%, **5%, ***1% significance.

	(1)	(2)
HPI growth measure:	3-year HPI growth from t-3 to t	2-year HPI growth from t-3 to t-1
2nd quartile HPI growth CBSAs	0.0105	0.0191
	(0.0131)	(0.0118)
3rd quartile HPI growth CBSAs	0.0322**	0.0412***
	(0.0159)	(0.0150)
4th quartile HPI growth CBSAs	0.0759***	0.0733***
	(0.0179)	(0.0169)
Annual average employment growth	-0.164	-0.113
	(0.142)	(0.140)
Average weekly wage growth	-0.0745	-0.0510
	(0.0912)	(0.0924)
Average quarterly growth in number of establishments	0.423**	0.450***
	(0.166)	(0.169)
Unemployment rate	0.418	0.264
	(0.568)	(0.571)
Year dummies	Yes	Yes
Age dummies	Yes	Yes
CBSA dummies	Yes	Yes
Observations	486,274	486,274

Table VI: Hazard Rate of Buying the First Home (Subsample for Individuals Who Enter the Sample with Credit Score $\geq$ 660)

The table estimates the hazard rate of buying the first home using a probit binary choice model for first home purchase. The sample is based on the 5% sample of the primary-individual data from the FRBNY (Equifax) Consumer Credit Panel Data set, and only includes individuals who enter the sample with credit score greater than or equal to 660. Individuals exit the sample after the year of their first home purchases. The main variable is the quartile dummies sorted on the HPI growth among the pooled CBSA-year observations. Control variables are the time-varying individual-level credit score, CBSA level employment and wage variables, year, age, and CBSA effects. Standard errors are clustered at CBSA level and are shown below the coefficient estimates. *10%, **5%, ***1% significance.

	(1)	(2)
HPI growth measure:	3-year HPI growth from $t-3$ to t	2-year HPI growth from $t-3$ to $t-1$
2nd quartile HPI growth CBSAs	0.0746***	0.0561**
	(0.0257)	(0.0244)
3rd quartile HPI growth CBSAs	0.0684**	0.0618**
	(0.0325)	(0.0287)
4th quartile HPI growth CBSAs	0.0916***	0.0817***
	(0.0324)	(0.0298)
Credit score (time-varying)	-0.000416***	-0.000415***
	(0.000121)	(0.000121)
Annual average employment growth	0.170	0.209
	(0.337)	(0.340)
Average weekly wage growth	0.591**	0.596*
	(0.302)	(0.305)
Average quarterly growth in number of establishments	0.126	0.107
	(0.248)	(0.252)
Unemployment rate	-0.914	-1.043
	(0.814)	(0.784)
Year dummies	Yes	Yes
Age dummies	Yes	Yes
CBSA dummies	Yes	Yes
Observations	209,009	209,009

Table VII: Quantile Regression of Loan-to-value (LTV) Ratio

The table reports the results of regressing the median and third quantile of loan-to-value (LTV) ratio on lagged housing price growth using quantile regressions. The sample only includes homebuyers of the 1% sample of the primary-individual data from FRBNY (Equifax) Consumer Credit Panel Dataset whose first time home purchase mortgage can be matched to McDash mortgage servicing data using the Credit Risk Insight Servicing McDash (CRISM) dataset. The main variables are 3-year lagged HPI growth and the quartile dummies sorted on the HPI growth among the pooled CBSA-year observations. Control variables are year and age effects. Standard errors are clustered at the CBSA level and are shown below the coefficient estimates. *10%, **5%, ***1% significance.

	(1)	(2)	(3)	(4)
	Quantile Regression			
	50th LTV		75th LTV	
2nd quartile HPI growth CBSAs		0.767		-0.0300
		(0.860)		(0.856)
3rd quartile HPI growth CBSAs		-1.365		-0.230
		(1.073)		(1.067)
4th quartile HPI growth CBSAs		-6.805***		-5.380***
		(1.044)		(1.039)
3-year HPI growth from t-3 to t	-23.78***		-29.19***	
	(3.823)		(3.726)	
Year dummies	Yes	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes	Yes
Observations	7,248	7,248	7,248	7,248

Table VIII: Housing size and HPI growth

The table reports the results of regressing housing size (proxied by housing price in 2000 dollars) on quartile dummies sorted on HPI growth among the pooled CBSA-year observations by age. Housing price in 2000 dollars is the appraisal amount (i.e. origination amount/LTV) deflated by CBSA-specific house price growth to 2000 dollars. The sample only includes homebuyers of the 1% sample of the primary-individual data from FRBNY (Equifax) Consumer Credit Panel Dataset whose first time home purchase mortgage can be matched to McDash mortgage servicing data using the Credit Risk Insight Servicing McDash (CRISM) dataset. Columns 1 and 2 reports the results of the regressions when the sample is restricted to homebuyers whose age was less than 30 at time of home purchase. Columns 3 and 4 reports the results of homebuyers whose age was greater than or equal to 30 at time of home purchase. The main variables are the quartile dummies sorted on the HPI growth among the pooled CBSA-year observations. Control variables are the individual-level credit score at time of home purchase, CBSA-level employment and wage variables, year, age, and CBSA effects. Standard errors are clustered at the CBSA level and are shown below the coefficient estimates. *10%, **5%, ***1% significance.

	(1)	(2)	(3)	(4)
Dependent variable:	Housing Size			
	Young	Young	Old	Old
2nd quartile HPI growth CBSAs	-14,473**	-16,039**	3,111	2,562
	(6,369)	(6,258)	(7,061)	(7,766)
3rd quartile HPI growth CBSAs	-27,675***	-29,764***	9,842	9,575
	(7,929)	(7,833)	(8,624)	(9,355)
4th quartile HPI growth CBSAs	-20,665***	-23,043***	20,541**	20,926*
	(7,826)	(8,264)	(9,713)	(11,102)
Credit score (time-varying)	223.3***	225.1***	341.3***	341.7***
	(26.30)	(26.84)	(34.50)	(34.61)
Annual average employment growth		-1,163		124,024
		(106,262)		(126,729)
Average weekly wage growth		40,893		119,530
		(122,084)		(155,888)
Average quarterly growth in number of establishments		-74,040		-112,479
		(92,697)		(79,744)
Unemployment rate		-191,941		110,373
		(294,008)		(231,691)
Constant	99,020	117,244	-162,070***	-173,117***
	(148,524)	(151,325)	(29,673)	(31,406)
Year dummies	Yes	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes	Yes
CBSA dummies	Yes	Yes	Yes	Yes
R-squared	0.332	0.332	0.303	0.303
Observations	2,807	2,805	4,508	4,506

Table IX: Downpayment and HPI growth

The table reports the results of regressing home purchase downpayment on quartile dummies sorted on HPI growth among the pooled CBSA-year observations by age. Downpayment is defined to be loan size $\times$(1-LTV)/LTV. The sample only includes homebuyers of the 1% sample of the primary-individual data from FRBNY (Equifax) Consumer Credit Panel Dataset whose first time home purchase mortgage can be matched to McDash mortgage servicing data using the Credit Risk Insight Servicing McDash (CRISM) dataset. Columns 1 and 2 reports the results of the regressions when the sample is restricted to homebuyers whose age was less than 30 at time of home purchase. Columns 3 and 4 reports the results of homebuyers whose age was greater than or equal to 30 at time of home purchase. The main variables are the quartile dummies sorted on the HPI growth among the pooled CBSA-year observations. Control variables are the individual-level credit score at time of home purchase, house price in 2000 dollars, CBSA-level employment and wage variables, year, age, and CBSA effects. Standard errors are clustered at the CBSA level and are shown below the coefficient estimates. *10%, **5%, ***1% significance.

	(1)	(2)	(3)	(4)
Dependent variable:	Downpayment			
	Young	Young	Old	Old
2nd quartile HPI growth CBSAs	14,118*	11,293	17,383**	12,521
	(7,533)	(6,912)	(7,942)	(7,700)
3rd quartile HPI growth CBSAs	22,697*	18,796*	23,037***	17,196**
	(12,084)	(11,170)	(7,681)	(7,500)
4th quartile HPI growth CBSAs	27,144***	20,988**	26,100***	16,857**
	(9,925)	(8,784)	(6,560)	(7,387)
House Price (2000 Dollars)	0.901***	0.899***	0.678***	0.678***
	(0.314)	(0.313)	(0.0476)	(0.0479)
Credit score (time-varying)	-36.38	-33.06	46.53***	47.12***
	(54.03)	(52.16)	(14.33)	(14.05)
Annual average employment growth		-126,063		-48,946
		(89,593)		(88,919)
Average weekly wage growth		22,647		-184,646**
		(97,011)		(93,374)
Average quarterly growth in number of establishments		-87,512		20,025
		(81,874)		(59,621)
Unemployment rate		-593,074**		-757,369***
		(242,264)		(156,631)
Constant	-7,971	45,800	-102,327***	-77,943***
	(51,040)	(38,306)	(16,778)	(14,511)
Year dummies	Yes	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes	Yes
CBSA dummies	Yes	Yes	Yes	Yes
R-squared	0.576	0.578	0.668	0.671
Observations	2,727	2,727	4,410	4,408

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

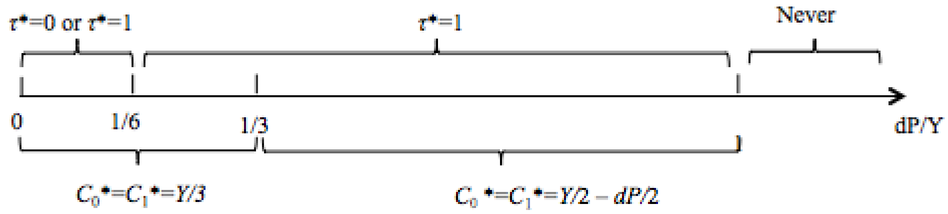


Figure A.2: Regions for Discussing Optimal Timing of First Home Purchases (Case 3)

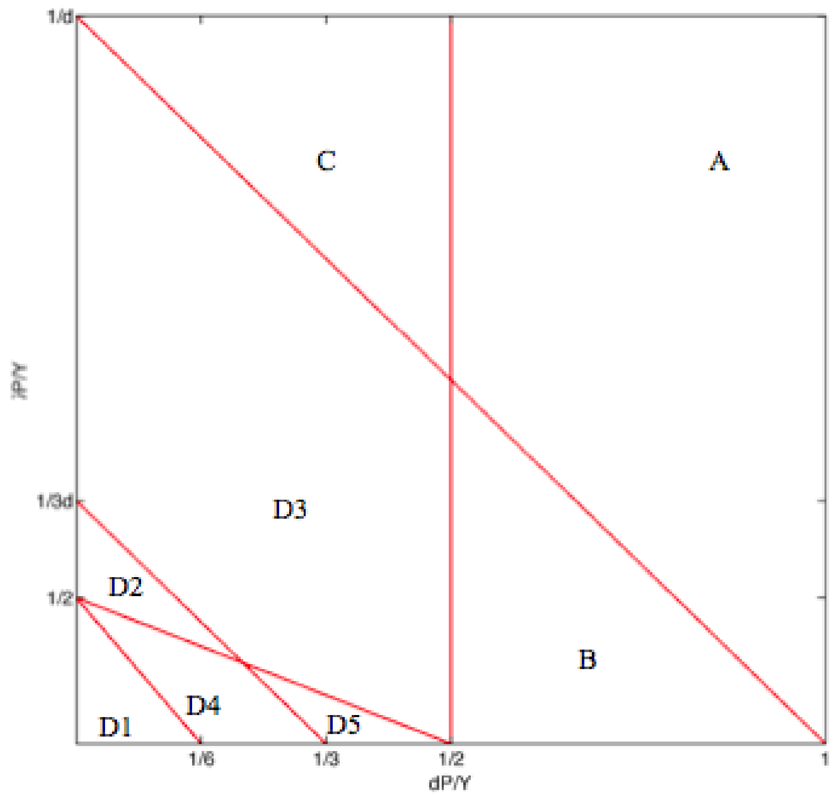
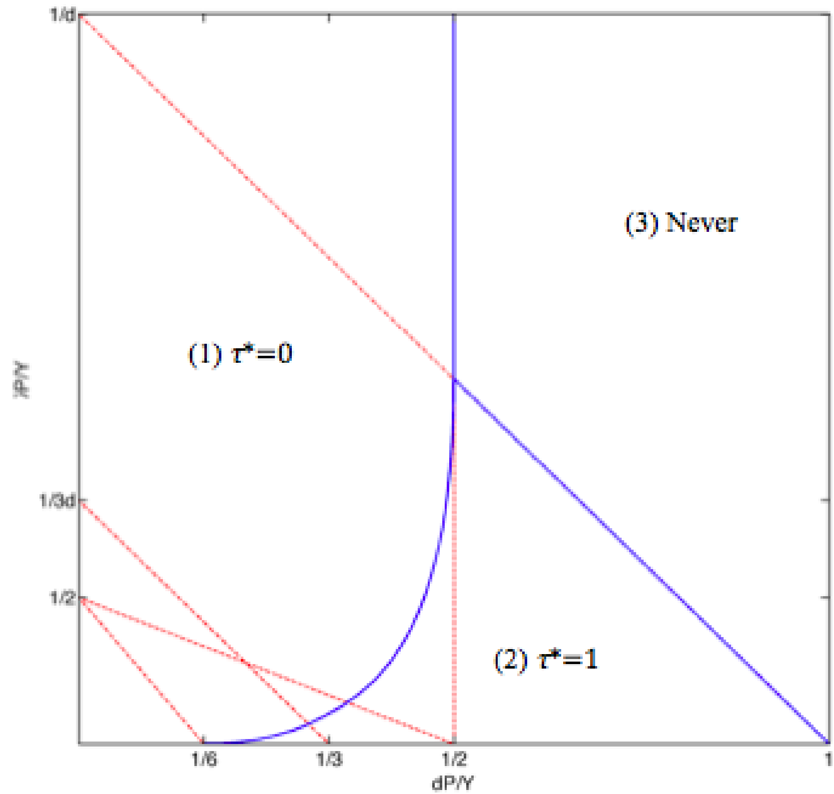


Figure A.3: Regions of Optimal Timing of First Home Purchases (Case 3)



Working Paper Series

A series of research studies on regional economic issues relating to the Seventh Federal Reserve District, and on financial and economic topics.

Working Paper Series *(continued)*

Working Paper Series *(continued)*

Working Paper Series *(continued)*

Working Paper Series *(continued)*

Financial Incentives and Educational Investment: The Impact of Performance-Based Scholarships on Student Time Use *Lisa Barrow and Cecilia Elena Rouse*	**WP-13-07**
The Global Welfare Impact of China: Trade Integration and Technological Change *Julian di Giovanni, Andrei A. Levchenko, and Jing Zhang*	**WP-13-08**
Structural Change in an Open Economy *Timothy Uy, Kei-Mu Yi, and Jing Zhang*	**WP-13-09**
The Global Labor Market Impact of Emerging Giants: a Quantitative Assessment *Andrei A. Levchenko and Jing Zhang*	**WP-13-10**
Size-Dependent Regulations, Firm Size Distribution, and Reallocation *François Gourio and Nicolas Roys*	**WP-13-11**
Modeling the Evolution of Expectations and Uncertainty in General Equilibrium *Francesco Bianchi and Leonardo Melosi*	**WP-13-12**
Rushing into the American Dream? House Prices Growth and the Timing of Homeownership *Sumit Agarwal, Luojia Hu, and Xing Huang*	**WP-13-13**