Title
Essays in Finance and Welfare

Author
Hacamo, Isaac Issa

Publication Date
2014

Peer reviewed|Thesis/dissertation

Essays in Finance and Welfare

by Isaac Issa Hacamo

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Business Administration in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Annette Vissing-Jørgensen, Co-chair

Professor Adair Morse, Co-chair

Professor Enrico Moretti

Fall 2014

Essays in Finance and Welfare

Isaac Issa Hacamo

Abstract

by Isaac Issa Hacamo

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor Annette Vissing-Jørgensen, Co-chair

Professor Adair Morse, Co-chair

This dissertation explores the relationship between finance and welfare, focusing on how the expansion of credit supply in the beginning of the 2000s, leading to the 2007 financial crisis, affected U.S. households. This dissertation helps advance the study of how financial policies affect households' well-being.

In the first chapter, I study the effect of access to credit on family structure. There is a large debate over the welfare effects of the early 2000s housing boom and bust. One potentially important welfare effect is the impact of mortgage credit expansion on family structure. Exploiting pre-housing boom variation on the distribution of old homeowners who live alone and are older than 65, I conduct within-county analysis with zip code level data to causally identify the effect of access to credit on fertility outcomes through a channel associated with a more efficient reallocation of the existing housing stock among households. I examine two other housing channels, house wealth gains and new construction, and show that the most relevant channel is the reallocation, which allows young households to access space by either moving to larger homes or achieving homeownership earlier in their life-cycle. A one standard deviation increase in reallocation leads to a 6.4% increase in fertility from 2000 to 2006. The same increase in house prices leads to only a 2.7% increase, and in new construction leads to a 1.5% decline in fertility from 2000 to 2006. I estimate that approximately 500,000 babies were born between 2000 and 2006 because of the reallocation channel.

In the second chapter, I study the effect of housing demand on house prices through an interest rate channel. In the last housing boom, strong house price growth only lasted until 2005. Why did house price growth slowed down in 2006? This chapter studies the effect of interest rate changes on housing demand at the end of a housing boom and the subsequent effect on house prices. I use three different proxies for housing demand, based on Google search data on search terms likely to be used during the process of purchasing a home, such as "remax", "construction", or "real estate". A one-standard deviation increase in the change in interest in purchasing a home from 2005 to 2006, measured by the Google search volume, leads to a 0.4 standard deviation increase in contemporaneous house price growth. To identify the interest rate channel, I first compute a household income threshold for each county, defined as the necessary income to afford an interest-rate-only mortgage on a county's average loan in 2005. I then exploit the slope of the county's income distribution around this affordability threshold to estimate the fraction of households that, after an increase in mortgage interest rates from 2005 to 2006, could no longer afford to pay an interest-rate-only mortgage. I use this fraction as an instrumental variable for housing demand. The IV beta is remarkably close to the OLS beta, and confirms the large effect of housing demand on house prices in 2006 through an interest rate channel. This chapter sheds light on the transition process between the housing boom in the 2000s and the subsequent financial crisis, and contributes to a better understanding of the impact of monetary policy on housing demand and house prices at the end of a housing boom.

Contents

List of Figures

List of Tables

Acknowledgements

I am deeply indebted to my dissertation committee members Annette Vissing-Jørgensen, Adair Morse, Atif Mian, and Enrico Morretti. They have been exceptional mentors and advisors. Adair and Annette's patience in helping me to think critically about the research process has been invaluable.

I thank all my fellow PhD program colleagues at Haas and in the Economics Department. I am particularly grateful to Brian Ayash, Andres Donangelo, Paulo Issler, Pedro Gardete, Gisela Rua, James McLoughlin, Hari Phatak, Nishanth Rajan, Javed Ahmed, Pratish Patel, Vladimir Asriyan, Victoria Vanasco, Mathew Botsch, Matteo Maggiori, Francesco D'Acunto, Michael Weber, Nirupama Kulkarni, and Raymond Leung for all their help and useful discussions. A special thank you for Tomás Reyes with whom I shared much of the last six years. Without the invaluable administrative help of Kim Guilfoyle and June Wong, the last two years would have been twice as challenging.

Much of my mental balance was kept healthy by my housemates at the three co-ops where I lived in Berkeley and Oakland: The Coop!, the Junkyard, and the Fish House. I thank all of them for their warm and delicious meals timely served at 8pm, great conversations, and unforgettable camaraderie. I am particularly grateful to Ryan Firestone, Wendy Morales, Dave Graham-Squire, Greg Gillette, Laura Espinoza, Siri Colom, Pegah Rahmanian, James Gill, Elif Ertekin, Asaph Widmer-Cooper, and Alexander Snipes for their invaluable friendship. Settlers of Catan was responsible for many pleasurable evenings, despite considerably decreasing my hours of sleep.

Last but not the least, I thank Sara, Fatma, and Angelique for their unconditional support and love. I believe Nizar would have been proud of this achievement. Angelique deserves a degree in psychological support for graduate students. Her patience and love are inspiring.

1 Chapter 1

The topic of access to finance and welfare has been studied in a number of dimensions. For example, studies have sought to quantify the impact of access to finance on welfare via its effects on intertemporal consumption smoothing (Jappelli and Pistaferri 2011; Gertler, Levine, and Moretti 2009), college enrollment (Levine and Rubinstein 2013), and job choices after graduation (Shu 2013). Others have studied the welfare impact of finance by documenting returns to finance jobs (Philippon and Reshef 2012; Kaplan and Rauh 2010), the elasticity of income with respect to financial output (Philippon and Reshef 2013), and borrower's behavior associated with distress finance (Melzer 2011; Karlan and Zinman 2010; Morse 2011).

In this chapter, I introduce a new channel whereby access to credit can offer welfare improvements, namely in fertility outcomes.[1] Given that space and children are likely to be strong complements, the increase in the availability of mortgage credit during the U.S. housing boom, which is associated with a large increase in homeownership and home transactions, could have had a large impact on households' decisions to have children. This is because demographers have suggested that the transition from renting to homeownership is associated with an increase in fertility, arguably because households have access to more space (Felson and Solaun 1975; Kulu and Vikat 2007; Mulder and Billari 2010; Strom 2010). It is then possible that sizable changes in the number of births might have occurred due to the expansion of mortgage credit, allowing me to plausibly identify a causal effect of access to credit on fertility decisions. In short, the contribution of this chapter is the identification and quantification of an effect of access to credit on fertility decisions through a channel associated with a more efficient reallocation of the existing housing stock among households, which creates access to space for young households who want to expand their families.

My identification relies on the ability to isolate a reallocation channel — associated with access to space — from other causes of fertility choices such as changes in household permanent income or changes in house wealth. I do this by laying out three channels by which the housing market could affect

[1] To make this argument, it must be assumed that fertility choices are, on average, welfare-improving. Beyond revealed preference demographers commonly link fertility and welfare, e.g., Thomson and Brandreth (1995); Kohler, Behrman, and Skytthe (2005); Margolis and Myrskyla (2011). I proceed under the assumption that having children is welfare-improving.

fertility: a wealth channel and two space channels. The house wealth channel helps households who are homeowners finance child rearing. Dettling and Kearney (2011) study the effect of house wealth on household fertility decisions; using Metropolitan Statistical Area level house prices from 1996 to 2006, they find that a $10,000 increase in house prices is associated with a 0.8% increase in fertility rates across homeowners (5%) and renters (-2.4%). The space channels, on the other hand, make it feasible to accommodate another house member in the dwelling, and are associated with access to larger homes or first-time homeownership. The two space channels through which access to credit impacts fertility are new construction and more efficient reallocation of the housing stock among households. My goal is to isolate the space channel associated with reallocation as a new causal channel between access to credit and fertility. To this end, I first estimate, after controlling for the observable determinants of fertility[2] and including county effects, the three housing channels in an ordinary least squares (OLS) framework. I proxy the intensity of reallocation with the change in per capita mortgage origination. However, some mortgage origination is not associated with reallocation. By choosing an appropriate instrument and using a two-stage least squares (2SLS) approach, I isolate the mortgage origination associated with reallocation and address endogeneity concerns from the OLS estimation.

Although the OLS estimates reveal that the reallocation is the relevant housing channel, the OLS estimates can be biased in both directions. For example, male permanent income shocks relax households' budget constraint allowing them to fund child rearing and simultaneously obtain more easily a mortgage loan. In this case, the OLS estimates are biased upwards if permanent income cannot be precisely controlled. Conversely, a shock to the female's level of education, or potential labor income, creates a negative bias, since the female's opportunity cost of child rearing increases, while the chance of qualifying for a mortgage loan increases. Therefore, to credibly identify the effect of access to credit on households' fertility decisions through a reallocation channel, I need an instrumental variable that correlates with fertility through the channel of interest — reallocation — and not through any other unobservable factor that drives fertility.

My empirical design is then defined by three features. First, I assume that

[2]Joseph Hotz et al. (1997) survey the fertility literature in developed economics and report the following variables as the most well identified determinants of fertility: income, unemployment, wealth, education, age structure, race, and ethnicity.

the whole U.S. economy experienced an outward shift on supply of credit led by relaxation of credit standards (Mian and Sufi 2009; Keys, Mukherjee, Seru, and Vig 2010). Second, to control for geographical differences between cities, especially differences in labor and housing markets that could confound the identification, I include county effects in all estimations, hence only the zip code level variation within-county is used for identification. Third, I use zip code level variation in the fraction of homeowners who are older than 65 and live alone in 2000, henceforth *old homeowners*, to generate exogenous variation in the supply of houses that could easily be subject to reallocation. The source of variation of the instrument relies on the underlying motives that old households have to exit their houses. During the housing boom, old homeowners exited their houses because they could monetize their home values, could not afford to pay increasing property taxes, or suffered from age-related health adversities such as death or disability. I claim that the exit due to monetization and increasing property taxes is driven by the global increase in house prices that was caused by the credit supply shock. Some old homeowners have a reservation price for their houses that credit constrained households can only pay when credit standards are loosened. Other old homeowners sell their houses and move out of their neighborhood when, due to increases in property assessments induced by the credit boom, property taxes rise to unaffordable levels relative to their income. The exit due to age-related health adversities is purely exogenous. Between 2000 and 2006 and within-county, the change in mortgage origination per capita is much larger in zip codes with high fraction of old homeowners relative to zip codes with low fraction of old homeowners, implying that the rank condition is met. Moreover, the increase in homeownership for young households (age<44) and decrease in homeownership for old households (age>65) is also larger in zip codes with high fraction of old homeowners. The instrument then isolates mortgage origination associated with reallocation of young households with old homeowners who live alone. By projecting the change in mortgage origination per capita on the instrument defined as old homeowners, the first stage will pick up mortgage origination that is associated with reallocation.[3] The exclusion restriction is guaranteed by the assumption that the global increase in house prices in the beginning of the 2000s was not driven by an unobserv-

[3]I assume that within-county houses are on average larger than apartments and thus suitable for young households to form and expand their families. I present anecdotal evidence in section 2.a that supports this assumption.

able determinant of fertility. Since the IV estimation controls for county effects, an alternative hypothesis has to drive all three variables - dependent, independent, and instrumental - in the average county in my sample. For example, if a permanent income shock is to confound the identification of the reallocation channel, it has to drive house prices and mortgage origination within the average county during the credit boom. However, Mian and Sufi (2009) show that between 2002 and 2005, and within-county, mortgage origination was disproportionally higher in zip codes with a high fraction of subprime borrowers despite their negative income growth.

One may still be concerned with the exclusion restriction of the aforementioned identification, particularly because unobservable income innovations could drive housing demand of credit constrained households and consequently cause the exit of old homeowners through monetization of high property taxes. Since the average life expectancy in the U.S. is 76 years for males and 81 years for females, I refine the above instrument by shifting the age limit to 75 years old, thus increasing the weight on the exit due to health-related reasons. The refined instrument is then the fraction of *homeowners who are older than 75 years old and live alone*, henceforth *75-homeowners*. Health adversities for people older than 75 are almost surely exogenous to possible unobservable income innovations that credit constrained households might have had during the credit boom. Although the nature of the instrument makes it unrelated to the credit shock, it generates exogenous variation in supply of housing that could be subject to reallocation during the housing boom. The first stage in the IV estimation in this second empirical exercise picks up more mortgage origination that is related with 'natural' reallocation. If the estimated coefficient is similar to the one estimated in the first empirical exercise, then it is plausible that the effect of credit supply induced reallocation on fertility is similar to the effect of 'natural' reallocation on fertility. I show that the two coefficients are indeed similar. Finally, one could be concerned that the reallocation identified during the credit boom happens any time, hence the estimated magnitude would be contaminated by the 'natural' shuffling between young and old households that constantly occurs in the economy. By examining the period from 1995 to 2000, right before the credit boom, I show that within-county the zip code level correlation between changes in mortgage origination per capita and changes in fertility is zero. The 'natural' reallocation in normal times is insufficient to create a correlation between mortgage origination and fertility — in opposition to what we observe during the housing boom.

To conduct my empirical analysis I construct a dataset of zip code level data that draws from a variety of data sources. I collect data on births from 10 Departments of Public Health: California, Idaho, Florida, Kansas, New York, Massachusetts, Oregon, South Carolina, Texas, and Wisconsin. I use individual loan data from Home Mortgage Disclosure Act (HMDA) to compute mortgage origination at the zip code level, and use income data from the Internal Revenue Service (IRS) to compute per capita income growth. I use data extracted from Zillow to compute zip code level house prices growth and use the Census and American Economic Survey to compute the demographic variables. The final dataset encompasses 2,753 zip codes, and covers approximately 70 million people in 2000, approximately 25% of the total U.S. population.

My estimates of the three housing channels show that during the housing boom the house wealth channel is not as large as estimated by Dettling and Kearney (2011). Using my zip code level dataset and the same regression specification as Dettling and Kearney (2011), I find that a $10,000 increase in house prices is only associated with a net annual increase of 0.4% in fertility rates, instead of 0.8%. One possible explanation for this difference could rely on the heterogeneity of house price growth across metropolitan areas between 1996 and 2006, since in contrast with the early 2000s, house price growth from 1996 to 2000 happened mainly in geographies with high income growth (Glaeser, Gottlieb, and Tobio 2012; Ferreira and Gyourko 2011). Dettling and Kearney (2011)'s results could be drawing from the beginning of the sample, while mine draw from the second part of the time period they analyze.

According to my estimation, the house wealth and the new construction channel have modest effects on fertility. A one standard deviation increase in house prices growth leads to a 2.7% increase in fertility from 2000 to 2006. When measured by zip code level growth in the number of housing units, the space effect due to new construction is negative. A one standard deviation increase in new construction leads to a 1.5% decline in fertility from 2000 to 2006. The negative sign suggests that new construction is associated with older households who have passed the fertility age.

By contrast, the reallocation of the existing housing stock has a larger impact; a one standard deviation increase in reallocation leads to a 6.4% increase in fertility from 2000 to 2006, which represents 28% of the standard deviation of fertility change. I then estimate the magnitude of the macroeconomic effect of the reallocation channel. For this purpose, I use county level

data since my zip code level dataset only covers 10 states and the county-level sensitivity is remarkably similar to the zip code level sensitivity. I start by sorting the counties by the change in the per capita mortgage origination change from 2000 to 2006. Next, I create 20 equal size bins, and, using the estimated coefficient from the IV regression, I estimate the change in fertility and number of births from 2000 to 2006 for each bin. I assume that the bottom bin is the control group, while the other bins are subject to the treatment effect; consequently, I subtract the number of births in the control bin from the treatment bins and sum the effects across all bins. Using this methodology, I estimate that 136,000 new births in 2006 are associated to the reallocation channel, corresponding to 3% of the total children born in 2006. If I assume that the increase in fertility is linear from 2001 to 2006, which I will argue is plausible as figure 1 suggests, then in 2001 the number of reallocation-related births is equal to 23,000, and the sum of all the births from 2001 to 2006 is equal to approximately 500,000 new births.

A child born during a pre-bust period could increase the pressure on households to seek additional disposable income during the bust. Households would be more likely to increase labor participation in an environment of high unemployment and provide less optimal early childhood education, which can ultimately affect future child outcomes. In an attempt to understand if such effects could be present, I conclude the chapter by presenting suggestive evidence that the change in fertility decisions due to the housing boom affected female labor participation during the financial crisis. Using individual records from the American Community Survey between 2007 and 2011, I show that women who had a child and lived in families who got a mortgage loan during the housing boom are more likely to be in the labor force during the financial crisis than similar women[4] who had a baby but rent in the same neighborhood (PUMA)[5]. They are however more likely to be unemployed than the average woman in the neighborhood, suggesting that they have a harder time obtaining a job because they stayed away from the labor force and chose to return during a time of high unemployment rates.

The remainder of this section presents the literature related to this chapter. The next section outlines the dataset used in this chapter, its construc-

[4]After controlling for head and wife's income, age, occupation, race, and education levels.

[5]Public Use Microdata Areas (PUMAs) are non-overlapping areas that partition each state into areas containing about 100,000 residents. PUMAs were developed to be the most detailed geographic area available in the Public Use Microdata Samples (PUMS).

tion, and summary statistics. Section 1.2 presents the empirical methodology, namely the housing-related mechanisms. Section 1.2 also lays out the empirical design to explore the causal effect of access to finance on fertility decisions through the reallocation channel. OLS and IV results are in the first part of section 2.3. The second part of section 2.3 presents robustness tests and the analysis of female participation in the labor force during the financial crisis. Finally, section 2.4 reports concluding remarks.

Related Literature. This chapter relates to three strands of literature. Firstly, it relates to the literature that studies the implications of the mortgage credit expansion and its welfare effects. Mian and Sufi (2009) and Keys, Mukherjee, Seru, and Vig (2010) seminal works show that in the beginning of the 2000s the U.S. economy experienced an outward shift in the supply of credit. Mian and Sufi (2009) document that less creditworthy borrowers experienced easier access to mortgage credit despite their negative income growth. Keys, Mukherjee, Seru, and Vig (2010) suggest that existing securitization practices adversely affected the screening incentives of subprime lenders. Adelino et al. (2012) use exogenous changes in the conforming loan limit as an instrument for lower cost of financing and higher supply to show that easier access to credit significantly increases house prices. Motivated by these findings and the severity of the financial crisis, a subsequent literature started examining the welfare effects of the expansion of credit and the role of finance in the past decades. Greenwood and Scharfstein (2013) show that, starting in 1980, fees associated with residential mortgages became a sizable portion of the growth in the U.S. financial services industry, while Philippon and Reshef (2012) show that workers in finance earned an education-adjusted wage premium of 50% in 2006, despite no premium in 1990. Charles, Hurst, and Notowidigdo (2013) suggest that housing booms disguise unemployment growth as they reduce the likelihood that displaced manufacturing workers remain unemployed. Mian and Sufi (2012) find that geographical differences in household debt overhang explain the differences of cross-sectional unemployment in the non-tradable sector. Levine and Rubinstein (2013) present evidence that intrastate bank deregulation increases the probability to attend college for individuals with particular learning abilities and family traits. Shu (2013) shows that careers in finance, especially at hedge funds and trading positions, attract students with high raw academic talent. This chapter adds to this literature by highlighting another welfare dimension that was affected by the expansion of mortgage credit - the family structure.

Secondly, the current chapter relates to the vast literature that studies the determinants of fertility. More than two centuries ago Malthus (1798) predicted a positive relation between income growth and population growth based on the hypothesis when people's incomes are higher they form families earlier and have more children. However, cross-national evidence over the last hundred years contradicts this prediction. As nations became industrialized and as their incomes increased, the fertility rate went down. Becker (1960), Becker and Lewis (1973) and Willis (1973) introduce the distinction between the quality and the number of children to explain the negative correlation between income and fertility. Angrist et al. (2010), however, show no evidence of a quantity-quality trade-off. Mincer (1963), Becker (1965), Willis (1973) and Schultz (1985) introduce women's time allocation decisions and emphasize the opportunity costs of women's time. Ermisch (1989) introduces market price of childcare to explain the impact of the mother's wage. Adsera (2005) suggests that the negative trend in fertility in developed countries is associated with constraints of the labor market where fertility decisions are taken. The cyclical behavior of fertility has received much attention since the work of Butz and Ward (1979). In most countries the fertility rate shows a negative response to unemployment along the business cycle, i.e., fertility is procyclical. Galor and Weil (1996) present a model where increases in women's wages lead to a decrease in fertility rates. Dettling and Kearney (2011) is the closest work to this chapter. They use MSA house price variation to study the effect of house wealth effect on fertility decisions from 1996 to 2006. This chapter reconciles their evidence with the other housing channels and highlights the importance of reallocation that stems from the relaxation of credit constraints.

Finally, this chapter relates to the literature that studies the impact of fertility on other economic outcomes. The relationship between labor supply and fertility has been long studied. Angrist and Evans (1998) find substantial effects of fertility decisions on parents' labor supply. They show that female labor supply effects appear to be absent among more educated women, while there is no relationship between wives' child-bearing and husbands' labor supply. Black et al. (2005) use a rich data set on the entire population of Norway and find a negative correlation between family size and children's education, but when they include indicators for birth order or use twin births as an instrument, family size effects become insignificant. Love (2009) presents a model in which marital status and children change savings behavior as well as portfolio choice. Bertocchi et al. (2011) study the join impact of

gender and marital status on financial investments by testing the hypothesis that marriage represents, in a portfolio framework, a safe asset. They show that married individuals have higher propensity to invest in risky assets than single ones.

1.1 Data

1.a Macroeconomic Indicators

Before discussing the micro dataset that I use to study the causal relationship between access to credit and fertility, I show that, in the last 20 years, the relationship between mortgage origination and fertility is present in the aggregate data only during the housing boom. The top panel in figure 1 shows that the aggregate number of births in the U.S. started an uptrend in 1996 that lasted until the end of the housing boom. The middle panel shows that, over the same time period, the fertility rate[6] exhibited an uptrend between 2000 and 2007. Both time-series suggest a shift in fertility choices during the housing boom period. The bottom panel of figure 1 shows that the annual volume of mortgage origination for home purchase shifted to a higher level between 2000 and 2006. Figure 2 confirms that households used mortgage loans to purchase existing and newly constructed houses by showing that the number of home transactions increased faster between 2000 and 2006. Figure 2 also shows that the number of transactions of existing houses was significantly larger than the number of newly constructed houses. This difference suggests that during the housing boom households were more likely to move into an existing house than a newly constructed one. Figure 3, using county-level data, proceeds to investigate the potential relationship between access to credit and fertility by showing that since 1995 mortgage origination and fertility are only positively correlated in changes between 2000 and 2006. The absence of correlation from 1995 to 2000, a period of strong economic growth, raises the bar for the permanent income hypothesis to be a credible alternative hypothesis. In order for permanent income to explain the positive correlation between fertility change and per capita mortgage origination change from 2000 to 2006 the correlation of income

[6] According to the CDC, fertility rate is defined as the number of births divided by the number of women in child bearing age, assumed to be from 15 to 44 years old.

growth and per capita mortgage origination change would need to change from 1995-2000 to 2000-2006.

Lastly, I sort the counties on the per capita mortgage origination change from 2000 to 2006 and depict in figure 4 the time series of the fertility rate for the top and bottom quintiles between 1990 and 2010. Prior to 1996, fertility rates are not statistically different between the two groups. By 2000, the difference is small; however, between 2000 and 2006, fertility rates increased rapidly in high mortgage origination counties; yet, in low mortgage origination counties fertility rates remained fairly constant. In sum, the macro evidence suggests that access to finance was strongly associated with fertility decisions during the credit boom.

1.b Micro Data

I draw from a variety of data sources to construct the sample used in this chapter. The sample consists of data on births, loans, income, house prices, employment, and demographics. Data on births is available by county and zip code, and was collected from the Department of Public Health (DPH) of each state. Birth statistics at the county level are available for 48 states from 2000 to 2006.[7]. Birth statistics at zip code level is available for 10 states: California, Idaho, Florida, Kansas, New York, Massachusetts, Oregon, South Carolina, Texas, and Wisconsin. For confidentiality reasons, in some states birth statistics are not available when the number of births is smaller than five in a given geography. Data at the zip code level is available for years 2000 and 2006.

Home Mortgage Disclosure Act (HMDA) provides loan level data from 1990 to 2011. Loan level data is publicly available for lenders that meet a disclosure criteria defined by HMDA every year. Each loan application provides information on *year* of application, *lender*, *type of loan*, *loan amount*, *action taken* by the lender, *reason for denial*, in case the loan is denied, *race*, *sex* and *income* of the applicant and co-applicant, *census tract*, *county FIPS*, and *state FIPS* where the loan was originated, *owner occupancy*, and *purpose*. Loans have four types of purpose: *home purchase*, *home improvement*, *refinancing*, and *multifamily dwelling*. I only use loans that are originated for home purchase and are owner-occupied as principal dwelling.

I use the Internal Revenue Service (IRS) data to compute the zip code

[7]The DPH of the state of Delaware and Louisiana did not make available their data at county level.

level income per capita. The IRS provides zip code level data for years 2001 and 2006. The provided income data includes adjusted gross income, number of returns, and wage income. Income per capita is defined as the ratio of the adjusted gross income to the number of returns.

Home prices are from Zillow. I extracted their sales-price-based price index for zip codes that have sufficient transaction level. Each Zillow Home Value Index (ZHVI) is a time series tracking the monthly median home value in a particular geographical region. In general, each ZHVI time series begins in April 1996. Instead of using a repeat sales methodology, Zillow uses the same underlying deed data as the Case-Shiller index but creates a hedonically adjusted price index. The Zillow index uses detailed information about the property, collected from public records, including the size of the house, the number of bedrooms, and the number of bathrooms. To the extent that the average measured characteristics of the home change over time, the Zillow index will capture such changes.[8] Guerrieri, Hartley, and Hurst (2013) show that the correlation between Case-Shiller Index and Zillow Index where the two samples overlap is equal to 94%. Monthly home prices are available from 1996 to 2012 for 10,187 zip codes.

Data on employment is from the County of Business Patterns (CBP) annual survey. CBP provides total employment for all establishments located in a given zip code. However, employment count from CBP is different from the employment for the zip code residents; therefore, I also use the employment data from the Decennial Census and the American Community Survey.

Finally, I use the public data from the Decennial Census and the American Economic Survey to obtain zip code data on gender, race, ethnicity, type of household, educational attainment, housing tenure, and number of bedrooms. The 2000 Decennial Census provides zip code data directly. On the other hand, to access the zip code data from the ACS, one needs to use 5-year averages. I use the ACS's 5-year averages from 2005 to 2009.

The construction of the dataset proceeds as follows: I start by merging the births and the Zillow Price data. The merged data set covers 3,256 zip codes. I proceed to merge it with HMDA data, and the number of merged

[8]More information about the computation methodology of the Zillow home price index can be found here: http://www.zillow.com/blog/research/2012/01/21/zillow-home-value-index-methodology/.

zip codes drops to 2,825. I then merge it with the IRS data and the CBP data, and as a result the number of zip codes drops to 2,793. Finally, after merging with the Census and ACS dataset the number of zip codes is 2,792. I then drop data points where births are missing in either 2000 or 2006, and repeat the same criteria for house prices and income data. The resulting dataset encompasses 2,753 zip codes, and covers 68.3 million people in 2000, approximately 25% of the total U.S. population.

Summary Statistics

Table 1 presents the summary statistics of the variables presented in this section. The change in per capita mortgage origination between 2000 and 2006 is equal to 4.7 per 1000 people on average, while the growth in house price during the same period is 106% in the average zip code, corresponding to annualized growth of 12.8%. For the 6 year period of analysis, the income growth is 22%, or in annualized terms 3.3%. Change in the female unemployment rate from 2000 to 2006 is 1.1%. Fertility changed by 3.37 per 1000 women in child bearing age from 2000 to 2006. The change in the zip code fraction of Hispanics and Blacks is on average 2.8% and 0.33%, respectively. The average population is 24,700 in 2000 and 26,700 in 2006. The average fraction of homeownership in 2000 is 65%. Finally, from 2000 to 2006 the average change in the fraction women with college degree is -1.7%.

1.2 Empirical Methodology

1.a Mechanism

If households are credit constrained, an outward shift in the supply of mortgage credit induces them to adjust their housing consumption. As a result, households move within the existing housing stock as well as to newly constructed houses. They access more space as renters become first-time homeowners and homeowners move into larger or better quality houses. I assume that the transition from renting to homeownership provides households additional housing space.[9] The credit supply shock then provides households

[9]I assume that the supply of apartments with more than two bedrooms is thin. Using 30 million Craigslist ads from 2008 to 2013, figure 5 presents suggestive evidence of thinness in the rental market. The price difference from a one to two bedroom is on average $390. By contrast, the price difference from a two to three bedroom apartment is $1100, and

access to space that would have been unaccessible otherwise or would have only been reachable later in their life-cycle. As households access more space, they presumably change their consumption of complementary goods; specifically, if they have a cobb-douglas utility function for housing and children, as I present in the stylized model in section 1.d, they may increase their fertility. A similar argument between space and fertility has been suggested by demographers (Felson and Solaun 1975; Kulu and Vikat 2007; Mulder and Billari 2010; Strom 2010). When a household moves into an existing house, I define it as reallocation. The credit supply shock can then affect the household's fertility decisions through two space channels: a reallocation and a new construction channel. Furthermore, since the credit shock causes an outward shift in housing demand it also impacts house prices. Increases in house prices create a wealth effect that relaxes homeowners' budget constraints, allowing better financing of child rearing, which increases the probability of having a child. The credit shock can then affect the household fertility decision through three channels: house wealth, new construction, and reallocation. The picture below outlines the three channels:

$$\begin{bmatrix}\text{Mortgage}\\ \text{Credit}\\ \text{Shock}\end{bmatrix} \rightarrow \begin{array}{c|l} \text{House Wealth} & \rightarrow \\ \hline \text{Space: Reallocation} & \rightarrow \\ \hline \text{Space: New Construction} & \rightarrow \end{array} \begin{bmatrix}\text{Family}\\ \text{Formation}\\ \text{and Expansion}\end{bmatrix}$$

Although reallocation could happen between any two households, in this chapter I will stress the reallocation between young and old households. The reasoning goes as follows: young households are typically credit constrained and more prone to family formation and expansion. By contrast, old households are more likely to reduce their housing consumption, especially old householders who own a house and live alone.

1.b Measurement

In my empirical analysis, I study three housing channels that can link the expansion of credit and fertility decisions: wealth gains from house price increases, new construction, and more efficient reallocation of the existing housing stock among households. In this section, I describe how I measure each channel as well as the outcome variable — the fertility change.

from a three to four bedroom is $2220. The higher relative increase in rent prices for larger apartments suggests short supply of large dwellings in the U.S. rental market.

House Wealth Channel

House prices impact family formation and expansion of homeowners and non-owners differently. An increase in house prices creates a wealth effect on homeowners, but tightens the budget constraint for renters since it increases the cost of housing. To distinguish the effect on homeowners and renters I interact the house price growth and the initial level of homeownership, since the house price effect is larger in zip codes where the level of homeownership is larger. The house price effect for a zip code i is:

$$\begin{aligned} \text{House Wealth Measure }_{2000\to2006,i} &= \lambda_1 \times \text{HP Growth }_{2000\to2006,i} \\ &\times \ \% \text{ Homeownership }_{2000,i} \\ &+ \ \lambda_2 \times \text{HP Growth }_{2000\to2006,i} \end{aligned}$$

In the regression specification I control for the level of homeownership in 2000.

Construction Channel

An exogenous shock in new construction affects house prices and housing consumption. House prices are affected through a pure supply channel. The effects on housing consumption depend on the relative size and quality of new houses constructed. To capture the space effect created by new construction, either from new housing units or houses remodeled, I will use the growth in the total number of bedrooms:

$$\text{Construction Measure }_{2000\to2006,i} = \frac{\#\text{House Units }_{2006,i} - \#\text{House Units }_{2000,i}}{\#\text{House Units }_{2000,i}}$$

Reallocation Channel

Credit constrained households have below optimal housing consumption. Inability to consume desired housing leads to lower demand and lower prices. A credit supply shock that lowers the lending standards facilitates the reallocation of housing resources. The ideal measure of reallocation quantifies the number of houses that are bought by credit constrained households from homeowners who were underutilizing their house. Since credit constrained households use the expansion of mortgage origination to buy a home, I proxy reallocation as the change in number of mortgages originated for home pur-

chase per household defined as:

$$\begin{aligned}\text{Reallocation Measure}_{\,2000\to2006,i} &= \left[\frac{\#\text{ Mortgage Origination}}{\#\text{ Population}}\right]_{2006,i} \\ &- \left[\frac{\#\text{ Mortgage Origination}}{\#\text{ Population}}\right]_{2000,i}\end{aligned}$$

Although Mian and Sufi (2009)'s work suggest that the above measure captures the increase in mortgage origination for credit constrained households, it is not guaranteed that the change in mortgage origination is correlated with purchases of larger homes. This issue is addressed by the instrumental variable.

Fertility Change

To measure fertility rates at the zip code level, I compute the ratio of the number of births over the number of women in child bearing age, assumed by the Centers for Disease Control and Prevention to be women with ages between 15 and 44. The Fertility change from 2000 to 2006 is then defined as:

$$\text{Fertility Change}_{\,00\to06,i} = \left[\frac{\#\text{ Births}}{\#\text{ Women}_{15<\text{age}<44}}\right]_{06,i} - \left[\frac{\#\text{ Births}}{\#\text{ Women}_{15<\text{age}<44}}\right]_{00,i}$$

1.c Estimation Methodology

My identification relies on the ability to isolate a space effect, associated with a better reallocation of the housing stock, from other causes of fertility choices - notably household permanent income. I do this by laying out three channels by which the housing market could affect fertility: a wealth channel and two space channels. The wealth effect helps households finance child rearing, but is only relevant to homeowners. The other housing channels that can explain fertility choices relate to space. Space makes it feasible to accommodate another house member in the dwelling and is provided by access to larger homes and first-time homeownership. The two channels by which space impacts fertility are new construction and efficient reallocation of the housing stock. My goal is to isolate the space channel, associated with reallocation, as a new causal relationship between access to finance and

fertility. I implement an empirical strategy in which the three effects are jointly estimated in an ordinary least squares framework. Then, using an instrumental variable approach, I isolate the space channel of interest and address issues related to endogeneity.

OLS

I first exploit zip code level variation with county effects to estimate the three housing effects that link access to finance with the change in fertility decisions. Since the regressors are likely to be endogenous, the estimates are potentially inconsistent. However, the direction of the coefficients and their magnitudes are informative about the potential economic significance of each channel. Furthermore, the comparison of the estimates with and without the other observable determinants of fertility is also informative about the stability of the coefficients and potential orthogonality of the effects with the error term. In all regression specifications, errors are robust and clustered at the state level. The regression model for zip code i is:

$$\begin{aligned} \text{Fertility Change}_{\,2000\to2006,i} &= \beta_0 + \beta_1 \times \text{Reallocation}_{\,2000\to2006,i} \\ &+ \beta_2 \times \text{Construction}_{\,2000\to2006,i} \\ &+ \beta_3 \times \text{House Wealth}_{\,2000\to2006,i} \\ &+ \alpha \times X_i + \text{County Effects} + \varepsilon_i. \end{aligned}$$

The literature on the economics of the family identifies fertility as a function of male and female income, wealth, unemployment, and the female's cost of time, race and ethnicity (Joseph Hotz et al. 1997, Butz and Ward 1979, Schultz 1985, Adsera 2005). Positive male permanent income shocks are associated with higher total fertility. More permanent income relaxes the household's budget constraint and allows it to finance child rearing. Transitory female income and unemployment shocks are associated with changes in fertility timing. Women time their fertility decision for times when their opportunity cost is small (Schultz 1985). Based on these identified *traditional determinants*, X_i includes controls for: per capita income growth form 2001 to 2006[10]; per capita log income in 2001; level of homeownership in 2000; per capita CBP employment change from 2000 to 2006; change in unemployment for women with ages between 25 and 44 from Census and ACS; change in

[10]IRS income data is not available at the zip code level in 2000, only in 2001.

composition of race and ethnicity; level of fertility in 2000; and change in fraction of college educated women with ages between 25 and 44. X_i also includes changes and levels of the fraction of women of ages from 15 to 17 years, 18 to 24 years, 25 to 34 years, and 35 to 44 years.

Endogeneity Concerns and IV The OLS estimates allow us to have a preliminary sense of what housing channels may have been relevant during the credit boom period. Although the OLS estimates suggest that the reallocation is the relevant channel to explain the effect of access to finance on fertility, they can be biased in both directions. For example, male permanent income shocks relax the household's budget constraints allowing it to fund child rearing and simultaneously more easily obtain a mortgage loan. In this case, if such a shocks are not appropriately controlled for, OLS estimates of the impact of reallocation on fertility is biased upwards. On the other hand, a shock to the female's level of education or potential labor income creates a negative bias since the female's opportunity cost of child rearing increases, but at the same time the chance of qualifying for a mortgage loan increase. To identify the effect of access to finance on the household's fertility decision, I need an instrumental variable that correlates with fertility through the channel of interest - reallocation - and not through any other unobservable factor that drives fertility.

The identification relies on three features. First, I assume that the whole U.S. economy experienced an outward shift on supply of credit led by relaxation of credit standards (Mian and Sufi 2009; Keys, Mukherjee, Seru, and Vig 2010). Second, I control for county effects to isolate geographical differences between cities, especially differences in labor and housing markets. Third, in order to instrument reallocation, I use zip code level variation in the fraction of homeowners who are older than 65 years old and live alone, henceforth *old homeowners*, and defined as:

$$\text{old homeowners} = \frac{\#\text{Homeowners than Older 65 and Living Alone }_{2000}}{\#\text{Households }_{2000}}.$$

During the housing boom, *old homeowners* exited their houses because they could monetize their home values, could not afford to pay increasing property taxes, or suffered from age-related health adversities such as death or disability. I claim that the exit due to monetization and property taxes is driven by the credit supply shock. Some *old homeowners* have a reservation

price for their houses that credit constrained households can only pay when credit standards are loosened. Other *old homeowners* sell their houses and move out of their neighborhood when, due to increases in property assessments induced by the credit boom, property taxes rise to unaffordable levels relative to their income. The exit due to age-related health adversities is purely exogenous. I assume that within the same county houses are on average larger than apartments[11] and thus suitable for young households to form and expand their families. The exit of *old homeowners* then generates an exogenous variation in the supply of houses available for reallocation that is unrelated to other determinants of fertility. Since the IV estimation controls for county effects, an alternative hypothesis has to drive all three variables - dependent, independent, and instrumental - in the average county in my sample. For a permanent income shock to confound the identification of the reallocation channel, it has to drive house prices and mortgage origination within the average county during the credit boom. However, Mian and Sufi (2009) show that between 2002 and 2005, and within-county, mortgage origination was disproportionally higher in zip codes with a high fraction of subprime borrowers despite their negative income growth.

One may still be concerned with the exclusion restriction of the aforementioned identification, particularly because unobservable income innovations could drive housing demand of credit constrained households and consequently cause the exit of *old homeowners* through monetization of high property taxes. Since the average life expectancy in the U.S. is 76 years for males and 81 years for females, I refine the above instrument by shifting the age limit to 75 years old, thus increasing the weight on the exit due to health-related reasons. The refined instrument is then the fraction of *homeowners who are older than 75 years old and live alone*, henceforth *75-homeowners*, and defined as:

$$\text{75-homeowners} = \frac{\#\text{Homeowners Older than 75 and Living Alone }_{2000}}{\#\text{Households }_{2000}}.$$

Health adversities for people older than 75 are almost surely exogenous to possible unobservable income innovations that credit constrained households might have had during the credit boom. Though the nature of the instrument makes it unrelated to the credit shock, it generates exogenous variation in supply of housing during the housing boom, and therefore it allows me to

[11] I present anecdotal evidence in section 2.a that supports this assumption.

identify the causal impact of access to finance (via reallocation channel) on fertility.

Lastly, one could raise doubts about the external validity of the results. Young credit constrained households can achieve space through reallocating by moving into to a vacant house, a house where the previous household was dissolved, a house where the current household upgrades to a larger house, or a house where the current household downsizes to a smaller house. Under these different options for reallocation, one should be concerned whether the treatment effect of reallocation on fertility produced by the variation of *old homeowners* is the same as the average treatment effect of all types of space increase on fertility (Angrist and Krueger 2001). The causal mechanism of reallocation is that access to space causes family expansion. Therefore, it seems plausible to assume that as long as young households access more space they will expand their families, despite who the previous homeowners were or the conditions that led the previous household to leave the house. Under this assumption, that space is the key variable, I assume that the local treatment effect estimated equals average treatment effect.

Table 3 panel A reports the OLS regression coefficients of *old homeowners* on the change in homeownership for different age groups, after controlling for county effects. Table 3 shows that the instrument captures precisely the variation of interest. In zip codes with high fraction of old homeowners, the homeownership increased for households whose head age is between 25 and 34 as well as between 35 and 44. The regression coefficient is equal to 0.21 and 0.28, respectively, and statistically significant. On the other hand, the regression coefficient on the change in homeownership for households whose head age is between 65 and 74 is -0.21, and above 75 is -0.29. The regression coefficient on the change in homeownership for households whose head age is between 44 and 65 is almost zero and insignificant. These regression coefficients show that, from 2000 to 2006, in zip codes with high fraction of *old homeowners* there was a reallocation of the housing stock whereby old households sold their houses to younger households. Table 4 panel A shows that the same results hold for the *75-homeowners*.

I finish this section presenting the regression model that implements the instrumental variable approach using the aforementioned instrument. I repeat the same specification when the instrument is refined from *old homeowners* to *75-homeowners*. I use a 2SLS estimation where in the *first stage*

I estimate:

$$\begin{aligned}
\text{Reallocation Measure}_{00\to06,i} &= \theta_0 + \theta_1 \\
&\times \left[\frac{\#\text{Owners, age>65 and Living Alone}}{\#\text{Households}}\right]_{00,i} \\
&+ \theta_2 \times \text{Construction Measure }_{2000\to2006,i} \\
&+ \theta_3 \times \text{House Price Measure }_{2000\to2006,i} \\
&+ \Theta \times X_i + \text{County Effects} + \eta_i,
\end{aligned}$$

and in the *second stage* I estimate:

$$\begin{aligned}
\text{Fertility Change }_{2000\to2006,i} &= \beta_0 + \beta_1 \times \widehat{\text{Reallocation Measure}}_{2000\to2006,i} \\
&+ \beta_2 \times \text{Construction Measure }_{2000\to2006,i} \\
&+ \beta_3 \times \text{House Price Measure }_{2000\to2006,i} \\
&+ \alpha \times X_i + \text{County Effects} + \varepsilon_i.
\end{aligned}$$

Vector X_i includes controls for: per capita income growth from 2001 to 2006, per capita log income in 2001, per capita employment change from County of Business Patterns, change in unemployment for women with ages between 25 and 44 from Census and ACS, change in composition of race and ethnicity, level of fertility in 2000, and change in fraction of college educated women with ages between 25 and 44. X_i also includes changes and levels of the fraction of women with ages from 15 to 17 years, 18 to 24 years, 25 to 34 years, and 35 to 44 years. The rationale for each control is presented above, in the OLS section.

1.d Stylized Model

The stylized model presented below formalizes the dynamics between housing consumption and demand for children when households are subject to a shock in access to credit. Credit constraints are modeled as a shock in downpayment requirements. Therefore, a shock in access to credit is modeled as a negative shock to downpayment requirements. Children are complements to housing. This simple framework delivers the empirical predictions of the mechanism described in 2.a. When downpayment payment requirements fall, households increase their housing consumption as well as their demand for children - the balance between timing and permanent change in fertility varies

with the model's parameters.

Consider a two period model where households consume housing, H, and children, C. Households have cobb-douglas preferences that makes them like more children when they have more housing. The price of housing consumption is given, equal to p, and priced in units of children consumption. Households choose their housing consumption at period $t = 0$, which is kept unchanged in $t = 1$, and choose the amount of children in each period separately. They can borrow at $t = 0$, but need to meet the downpayment constraint of γ times the dollar amount of housing services. The interest rate is r and the mortgage that was originated at $t = 0$ needs to be paid in full at $t = 1$. Households earn a constant wage w. They maximize:

$$U(H, C_t) = \sum_{t=0}^{1} \beta^t \left(\alpha \log(H) + (1-\alpha)\log(C_t)\right)$$

subject to:

$$\begin{aligned} \gamma p H + C_0 &\leq \omega \text{ at t=0, and} \\ (1-\gamma) p H (1+r) + C_1 &\leq \omega \text{ at t=1.} \end{aligned}$$

One can then show that under certain parameter conditions, C_0 is decreasing in γ. Particularly, when $1 + r = 1$ and $\beta = 1$ the closed form solution for derivative of C_0 with respect to γ is:

$$\frac{\partial C_0(\gamma)}{\partial \gamma} = \frac{2\alpha(2\gamma - 1)\omega}{(\gamma - 1)g(\gamma, \alpha)} - \omega \frac{1 + \alpha + g(\gamma, \alpha)}{4(\gamma - 1)^2}$$

where $g(\gamma, \alpha) = \sqrt{1 + \gamma^2 + 2\alpha(1 + 8(\gamma - 1)\gamma)}$. In this particular case C_0 decreases in γ when $\gamma < 1/2$. In other words, as long as the downpayment is smaller than 50% of the house price, households will choose to have more children early if they experience a negative shock in the downpayment requirements. Three predictions come out of the model when credit standards are relaxed for credit constrained households.

Prediction 1: $\partial H/\partial \gamma < 0$. Households demand more housing consumption when downpayment requirements decrease. This is the part of the model that delivers the housing boom. Downpayment requirements are one way by which credit standards were loosened. Relaxation of credit score require-

ments was likely the most common one. However, in the two period setting it is fairly tractable to implement downpayment requirements as it is above.

Prediction 2: $\partial C_0/\partial \gamma < 0$. Households demand more children in period $t = 0$ as credit standards are loosened. This is a feature of the cobb-douglas utility function, since the more housing households consume the more children they prefer.

Prediction 3: Children consumption in period $t = 1$ varies with the parameter choice, but total consumption of children increases, $\partial C_T/\partial \gamma < 0$, where $C_T = C_0 + C_1$ is the total number of children that the household decides to have in both periods. Since total fertility, C_T, increases while C_1 might decrease, some of the increase in fertility during the housing boom is a matter of timing. Young households, because of easy access to credit, obtain homeownership earlier in their life-cycle and consequently form their families earlier as well.

1.3 Results

Before reporting the main results of the chapter at the zip code level, I show that the relation between changes in fertility and changes in mortgage origination per capita is similar at the county and zip code level. The similarity between zip code level and county level results is important to estimate the aggregate effect of access to credit on fertility changes. The zip code sample only covers 25% of the U.S. population, while the county level sample covers approximately 93% of the population.

1.a OLS: County Level

Table 5 reports the OLS county level regression coefficients of the fertility rate change from 2000 to 2006 on traditional determinants and the housing channels.[12] From the baseline specification, without state effects and population weights, in column 1, the addition of the change in mortgage origination increases the R^2 from 13.7% to 19.9%, column 3. At the county

[12]To account for the mechanical changes in fertility rates that are caused by demographical migration between counties, I control for changes in and levels of age, race and ethnicity groups.

level, mortgage origination explains an additional 45% of the variation relative to traditional determinants. The inclusion of state effects decreases the coefficient on mortgage origination from 0.41 to 0.35 in the model with controls, and the inclusion of population weights changes the coefficient on mortgage origination to 0.30. Lastly, column 8 and 10 provide interesting county-level evidence about the three housing channels. First, the coefficient on mortgage origination (0.30) does not change after including the measures for the house wealth and new construction channel. Second, and surprisingly, the net effect of the house price growth is negative (-2.11) at the mean of U.S. homeownership (0.66). For a 10% increase in house price growth the number of births decreases by -0.211 per 1000 women in child bearing age from 2000 to 2006. Moreover, the magnitude is unchanged with the addition of mortgage origination in the regression model. This specification, since it includes state effects, differs from Dettling and Kearney (2011), who studied the effect of house price changes on fertility. The model with state effects suggests that the negative effect on fertility associated to renters outweighs the positive effect from wealth gains of homeowners. By contrast, the effect of per capita mortgage origination change on fertility change is equal to 0.30 and stable across specifications.

1.b OLS: Zipcode Level

Traditional Determinants

The previous literature has identified various determinants that explain fertility, namely male and female income, wealth, unemployment, and the female's cost of time, race and ethnicity (Joseph Hotz, Klerman, and Willis 1997, Butz and Ward 1979, Schultz 1985, Adsera 2005). Since I can measure these determinants, I start by estimating them and comparing the signs and magnitudes to the ones previously found in the literature. Table 6 reports that the change in unemployment of males who live in the zip code correlates negatively with the change in fertility and is statistically significant. On the other hand, female unemployment is positively correlated with the fertility change, although not statistically significant. Male unemployment reduces household's fertility because of the large negative income shock and uncertainty associated with the loss of employment (Butz and Ward 1979). Given male unemployment, female unemployment is commonly associated with timing whereby women choose to have children when their opportunity

cost of child rearing is low (Schultz 1985). The signs of the two unemployment coefficients match the ones found previously in the fertility literature. The growth in employment from the CBP survey measures the growth in employment from the businesses located in the zip code. CBP employment growth is positively correlated with fertility changes, capturing zip code level economic development, which potentially provides employment opportunities for residents. The level of income, measured by IRS data, is negatively correlated with fertility changes. Lower income households are more likely to have more children (Galor and Weil 2000); for example, teenagers in low income families tend to have higher fertility rates than teenagers in high income families. The per capita income growth effect, measured with the IRS data, on fertility is positive, but not statistically significant. The lack of significance is likely to be related to either the strong significance of the employment growth, or to the fact that income levels have strong prediction power of the income growth rates. Finally, zip code changes in females' college level education, which proxies for the change in the opportunity cost of the average female in the zip code, is negatively correlated with changes in fertility, consistent with the literature that claims that as females's opportunity cost increases fertility decreases (Schultz 1985). Table 7 reports the coefficients on the age and demographic variables of the same regression. Consistent with the literature on fertility (Parrado and Morgan 2008), zip codes where the fraction of Hispanics increases experience an increase in fertility. Likewise, if the zip code experiences an increase in the fraction of blacks it also experiences an increase in fertility.

1.c Discussion of the Housing Channels

I examine three housing channels that can link the expansion of credit and fertility decisions: wealth gains from house price increases, new construction, and more efficient reallocation of the existing housing stock among households. My goal is to isolate the space channel associated with reallocation as a new causal channel of access to credit on fertility. To that end, I first estimate, in an ordinary least squares framework, the three housing channels controlling for the traditional determinants of fertility. Then, using an instrumental variable approach, I isolate the mortgage origination associated with the reallocation channel and address endogeneity concerns that arise from the OLS estimation. I present the results of the IV estimation in the next section, while in this one, I discuss the OLS estimation based on the

results reported in table 6. The OLS estimation allows the following three inferences.

First, I find that the house wealth channel during the credit boom is not as large as estimated by Dettling and Kearney (2011). Dettling and Kearney (2011) use MSA level house price variation to study the house wealth effect on fertility decisions from 1996 to 2006. They find that a $10,000 increase in house prices is associated with a 5% increase in births among homeowners and a 2.4% decrease among non-owners. At the mean of U.S. homeownership rate the net effect is 0.8%. Using my zip code level dataset and the same regression specification as Dettling and Kearney (2011), I find that a $10,000 increase in house prices is only associated with an annual increase of 0.4% in births, instead of 0.8%. One possible explanation for this difference relies on the heterogeneity of house price growth across metropolitan areas between 1996 and 2006, since in contrast with the early 2000s, house price growth from 1996 to 2000 happened mainly in high income growth areas (Glaeser, Gottlieb, and Tobio 2012; Ferreira and Gyourko 2011). Dettling and Kearney (2011)'s results could be drawing from the beginning of their sample, while mine draw from the second part of the time period they analyze. In the OLS framework, I estimate that, at the mean of the U.S. homeownership (0.66), a one standard deviation increase in house prices growth (0.54) leads to a 2.7% increase in the fertility from 2000 to 2006, which is 13% of a standard deviation increase in fertility change.

My second inference is that, when measured by zip code growth in the number of bedrooms, the space effect due to new construction channel is negative. A one standard deviation increase in growth in number of bedrooms (0.23) leads to a 1.5% decline in fertility from 2000 to 2006, which is a 6% of a standard deviation decrease in fertility change. The negative sign suggests that new construction is associated with older households who have passed the fertility age. Moreover, the negative sign corroborates an equilibrium where households move up, meaning mid-age households move up to new houses and young households move into existing houses. Under this view, the new construction channel is consistent with the reallocation channel.

Third, reallocation, as measured by the change in per capita mortgage origination after controlling for the other housing channels, correlates positively with the change in fertility. The coefficient without controls (0.29) is remarkably close to the one estimated at the county-level. After accounting for other traditional determinants the OLS coefficient decreases to 0.20. Since the coefficient can be biased in both directions, I proceed to instru-

ment the variation of mortgage origination with the fraction of homeowners who live alone and are older than 65 (*old-homeowners*). The validity of the instrument is discussed in section 1.c.

1.d Instrumental Variable: old Homeowners

To eliminate the possible reallocation coefficient's bias in the OLS estimation, I use the fraction of homeowners who live alone and are older than 65, *old homeowners*, to generate variation in reallocation that is uncorrelated with other unobservable determinants of fertility rates. I discuss in section 1.c why the instrument is correlated with the measure of reallocation and why, conditional on controlling for reallocation, is likely to be uncorrelated with the other unobservable determinants of fertility.

In this section, I report the first-stage and the IV estimates of the second stage. Table 6 reports the first-stage together with the second-stage. In the first-stage, the estimated coefficient on old homeowners is 22.24 with a standard error of 4.09 and a F-statistic of 16.31.[13] The IV coefficient is 0.31 and statistically significant, with a t-statistic of 2.38. The instrumented reallocation measure implies that one standard deviation change in per capita mortgage origination change (12.0) leads to a 6.4% increase in fertility from 2000 to 2006, which is 28% of a standard deviation increase in fertility change. The estimated coefficients in the other two housing channels are almost unchanged, as well as the coefficients on the traditional determinants of fertility. The stability of the housing channels in the IV estimation relative to the OLS estimation makes it likely that the the reallocation measure is orthogonal to the house wealth and new construction measures. The IV coefficient (0.31) on the reallocation measure is almost equal to the OLS coefficient without controls (0.29), but 55% higher than the OLS with controls (0.20). Although measurement error is fairly plausible, an alternative explanation relies on the nature of the treatment associated with the instrument. The measure of reallocation, per capita mortgage origination, measures mortgage loans that were originated without the purpose of reallocation, for example, a mortgage originated to a single householder. Although the same could be claimed about the household who buys a house from a old homeowner, the estimated IV results suggest that it was more likely that the old homeowner was replaced by a couple who had a child around the time of purchase.

[13] According to Staiger and Stock (1997), who formalized the definition of weak instruments, since the F-statistic exceeds 10, the instrument is sufficiently strong.

The IV results point out the importance of the reallocation channel and the lack of power of the other two housing channels to explain the variation of the fertility between 2000 and 2006. Since the house wealth has gained relevance in the fertility literature, I investigate furthermore the impact of this channel on fertility. Table 8 shows that the net effect of the house wealth on fertility is even weaker than my IV estimates if I change the regression specification. If one wants to estimate the net effect of the house prices on fertility the inclusion of only house price should suffice to capture the net effect. In column 7 of Table 8, I present the estimates of the OLS regression in which the reallocation channel and new contraction are included together with the house price growth. The estimated coefficient on house price growth is significant at the 5% level and equal to 2.10, implying that a one standard deviation increase in house prices growth (0.54) leads to a 1.8% increase in the fertility from 2000 to 2006, which is 7.9% of a standard deviation increase in fertility change. Moreover, in the IV estimations the coefficient drops to 1.77 and the standard errors increase. Together, the evidence for the house wealth channel, suggests that between 2000 and 2006 the net impact is small. The increase in fertility of homeowners due to house wealth gains is cancelled out by the decrease in fertility of renters due to higher costs of housing.

1.e Refinement: 75-Homeowners

When the instrument is refined to focus only on the older (age>75) population that is more likely to exit their house because of health related reasons rather than the monetization and the price out, the F-statistic of the first-stage equals 16.74. The instrumented coefficient is 0.31 and statistically significant with a t-statistic of 2.2. When the instrument relies more heavily on the the clearly exogenous part of the variation in reallocation the coefficient is still estimated with the same magnitude. This evidence, as discussed in section 1.c, should alleviate concerns that the identification of the reallocation channel, presented in the previous section, is confounded by unobservable innovations in permanent income.

1.f Economic Magnitude

As expansion of mortgage credit is associated with a large increase in homeownership and home reallocations, it is plausible that sizable changes in the number of births occurred due to the housing boom. To estimate the magnitude of the effect of the access to credit through the reallocation channel, I first sort the counties by change in the per capita mortgage origination

from 2000 to 2006. I use the whole sample of counties where I have data on births and loans from HMDA, that is 2091 counties that cover approximately 93% of the U.S. population. I use county level data since my zip code level dataset only covers 10 states and the county-level sensitivity of fertility to the reallocation channel is remarkably similar to the zip code level IV sensitivity. I create 20 equal size bins, and using the estimated coefficient from the IV regression, I estimate the change in fertility for each bin from 2000 to 2006. Then, using the change in fertility and number of women of child bearing age in each bin, I compute the number of births in 2006 in each bin due to the reallocation channel. Finally, I assume the bottom bin to be the 'control' bin and the others to be the 'treatment' bins. The estimate of births in 2006 is then equal to the sum of 'treatment' bins minus 'control' bin. Using this methodology and relying on the assumption that the bottom bin is a fair 'control' group, the estimated number of births is equal to 136,000 in 2006. If I assume that the growth in fertility is linear from 2001 to 2006, as figure 1 suggests, then in 2001 I estimate the number of births to be 22,800. The sum of all the reallocation-related births from 2001 to 2006 is equal to 478,000. About 3% of the children that were born in 2006 were due to the housing boom.

1.g Female Participation in the Labor Force

According to the U.S. Department of Agriculture, it will cost, in 2012 dollars, approximately $248,000 for a middle-income family to raise a child for 18 years. A child born during a credit boom period could increase the pressure on a household to seek additional disposable income during the bust. The household would then be likely to increase their labor supply in an environment of high unemployment and potentially provide less optimal early childhood education to their children, which could ultimately affect future outcomes of their children.[14] According to this hypothesis, and assuming that the cost of owning is higher than renting, households who decide to have a child and get a mortgage loan during the housing boom period (*mortgage-baby* households) are more likely to experience financial distress during the crisis period than similar households who decided to be renters (*renter-baby* households). To test if women in mortgage-baby households are more likely

[14]Leventhal and Newman (2010) suggest that residential mobility during the 2008 financial crisis is associated with deleterious effects on children's short-term academic performance.

to be in the labor force during the bust, I use individual records from the American Community Survey. The American Community Survey surveys 3% U.S. households every year after 2005. The panel is cross-sectional and has no longitudinal dimension. However, ACS asks when the households move into their dwellings, as well as the type of housing tenure; and, in case of ownership, they ask if there is a mortgage loan. With this information, and assuming that the current mortgage existed since they moved in, I can back out the households who obtained a mortgage loan between 2000 and 2006, and compare them with similar households who moved into an apartment in the same neighborhood (PUMA).[15] The regression model for household h is:

$$\begin{aligned} \mathrm{Y}_h &= \beta_0 + \beta_1 \times \mathbb{1}_{\{\textit{Mortgage and Baby},h\}} \\ &+ \Gamma \times X_h + \text{PUMA Effects} + \text{Race Effects} + \varepsilon_h. \end{aligned}$$

where,

$$\mathbb{1}_{\{\text{Mortgage and Baby}\}} = \begin{cases} 1, \text{ if mortgage and baby between 2000 and 2006} \\ 0, \text{ if renter and baby between 2000 and 2006} \end{cases}$$

and Y_h is 1 if the female in the household is in the labor force and 0 otherwise. The controls include wife and husband's age, college indicator, and race. I also control for the number of bedrooms and the dwelling's year of construction. Since households decided to buy a house or be renters, there is a fair amount of selection bias in the above regression. Although I try to address the bias by controlling for observable characteristics that can reduce the selection bias, we should bear in mind the possible selection bias when interpreting the regression results. Table 14 shows the results of the aforementioned regression model. Between 2007 and 2011, females in mortgage-baby households are 14% more likely to be in the labor force than females in renter-baby households. The second column of table 14 presents the same regression model, but with two differences. First, the left-hand-side variable is 1 if the female is employed and 0 otherwise. Second, the right-hand-side

[15]Public Use Microdata Areas (PUMAs) are non-overlapping areas that partition each state into areas containing about 100,000 residents. PUMAs were developed to be the most detailed geographic area available in the Public Use Microdata Samples (PUMS).

variable is now:

$$\mathbb{1}_{\{\text{Mortgage and Baby}\}} = \begin{cases} 1, \text{ if mortgage and baby between 2000 and 2006} \\ 0, \text{ otherwise.} \end{cases}$$

Intuitively, this exercise compares the female in the *mortgage-baby* household with all other similar females in the neighborhood. The estimated coefficient shows that the female in the mortgage-baby household is more likely to be unemployed, suggesting that they have a harder time obtaining a job likely because they stayed away from the labor force and chose to return during a time of high unemployment.

1.4 Concluding Remarks

This chapter introduces a new welfare effect of access to finance whereby access to credit can offer welfare improvements, namely in fertility outcomes. I conduct within-county analysis with zip code level data to document that changes in mortgage origination are strongly associated with changes in fertility rates beyond traditional fertility determinants such as income and unemployment. I examine three housing channels that could explain this correlation: wealth gains from house price increases, new construction, and more efficient reallocation of the existing housing stock among households. I claim that after controlling for the house wealth and construction channel, mortgage origination measures the reallocation channel. The reallocation allows young households to move to larger homes or achieve homeownership earlier in their life-cycle, while older households can downsize their housing consumption. I exploit zip code level variation in fraction of homeowners older than 65 and living alone to causally identify the reallocation channel. During the housing boom, old homeowners exited their houses because they could monetize their home value, could not afford to pay increasing property taxes, or suffered from age-related health adversities such as death or disability. I claim that the exit due to monetization and property taxes is driven by the credit supply shock. Some old homeowners have a reservation price for their house that credit constrained households can only pay when credit standards are loosened. Other old homeowners sell their houses and move out of the neighborhood because property taxes raise to unaffordable levels relative to their income, when property assessments increase induced by the credit boom. The exit due to age-related health adversities is purely

exogenous. The variation generated by the instrument allows me estimate the causal effect of access to finance on fertility decisions through the reallocation channel. The IV estimates show that one standard deviation increase in reallocation leads to a 6.4% increase in fertility from 2000 to 2006, which represents 28% of the standard deviation of fertility change. The same increase in house prices leads to only a 2.7% increase in fertility from 2000 to 2006, and in new construction leads to a 1.5% decline in fertility from 2000 to 2006.

Such a large number of births could significantly affect other economic outcomes. According to the U.S. Department of Agriculture, it will cost, in 2012 dollars, approximately $248,000 for a middle-income family to raise a child for 18 years. A child born during a credit boom period could then increase the pressure on a household to seek additional disposable income during the bust that follows the boom. The household would then be likely to increase their labor supply in an environment of high unemployment and potentially provide less optimal early childhood education to their children, which could ultimately affect future outcomes of their children. I present suggestive evidence that the change in fertility decisions due to the housing boom affected female labor participation during the financial crisis. Using individual records from the American Community Survey between 2007 and 2011, I show that women who had a child and lived in families who got a mortgage loan during the housing boom are more likely to be in the labor force and unemployed during the financial crisis than similar women who had a baby but rent in the same neighborhood. Beyond the direct impact on utility and the impact on expenditures, fertility decisions produce significant changes at the aggregate level by affecting population growth and economic growth (Barro and Becker 1989 and Becker et al. 1990). Therefore, if the expansion of credit affected the fertility rate of U.S. households, it is relevant to estimate the magnitude of the aggregate effect. I estimate that approximately 500,000 babies were born between 2000 and 2006. In 2006 3% of the total number of births was due to the reallocation channel. This chapter not only contributes to the literature on the welfare effects associated to the access to finance, but also points out a new determinant of fertility that was previously unidentified.

2 Chapter 2

In September 2008, when Lehman Brothers filed for bankruptcy, house prices had already been declining for almost three years in some counties. The extraordinary growth in house prices in the beginning of the 2000s lasted until the end of 2005. During 2006, the U.S. average house price leveled off. However, across cities, house price growth in 2006 was remarkably different. For instance, in Baltimore City, MD, house prices grew by 19.2% during 2006, while in Santa Barbara County, CA, they dropped by 11.3%. What caused this heterogeneity in house prices across cities in 2006? Was it related with the contractionary monetary policy that started in the summer of 2004? Was this cross-sectional variation associated with the house price decline that occurred during the financial crisis?

In this chapter, I study the effect of interest rate changes on housing demand at the end of a housing boom and the subsequent effect on house prices. This chapter sheds light on the transition process between the housing boom in the 2000s and the subsequent financial crisis. During the beginning of the 2000s, while the U.S. economy experienced a housing boom, house price growth reached historical highs, mortgage credit expanded, and many U.S. households with poor credit history accessed the mortgage credit markets (Mian and Sufi 2009, Keys, Mukherjee, Seru, and Vig 2010, Adelino, Schoar, and Severino 2012, Agarwal and Ben-David 2014). Since interest rates stayed at low levels from 2001 until the end of 2004, mortgage products that were sensitive to interest rates, such as interest-rate-only and adjustable-rate mortgages, became popular and gained a large market share (Mayer, Pence, and Sherlund 2009). Economic growth also followed the housing boom, and in the summer 2004, the Federal Reserve Bank slowly started to implement a contractionary monetary policy. In the beginning of the 2005 the federal funds rate interest rates crossed the 2% threshold. By the middle of 2006, the federal funds rate was at 5.25%. Therefore, from the beginning of 2005 until the middle of 2006, the one-year adjustable-rate increased by approximately 1.5%. Such a sudden and large increase must have shifted the housing demand curve inward, particularly for households who were close to their affordability thresholds at current house prices. The hypothesis of this chapter is that the increase in interest rates from 2005 to 2006 lowered the households's interest in purchasing a home, shifting the housing demand curve inward, which then created a downward pressure in house prices from 2005

to 2006.

Google searches have been shown to forecast short-term economic outcomes such as automobile sales, unemployment claims, or travel destination planning (Choi and Varian 2012). Vlastakis and Markellos (2012) use Google searches to measure information demand. Goel, Hofman, Lahaie, Pennock, and Watts (2010) show that what consumers search for online can predict their future behavior days or even weeks in advance. Da, Engelberg, and Gao (2011) use Google searches to measure investor attention and show that an increase in Google searches predicts higher stock prices. In this chapter, I proxy housing demand using Google's search volume index (SVI) on terms that are searched for during the process of purchasing a home. I build three different proxies. The first uses search terms associated with real estate firms, such as "remax", "century 21", "coldwell", and "zillow". The second measure uses general terms associated with real estate, such as "real estate", "realtor", or "mls". The third measure uses search terms associated with housing in general, such as "houses", or "construction". These search terms are present in queries such as "real estate market in Los Angeles", "mls listings", or "new contraction in Phoenix". My assumption is that, the more households search these terms on Google, the more interested they are in purchasing a home. The time series of these three measures is consistent with the timing of the housing boom and crisis. Furthermore, the change in 2006 in search volume at the county level is strongly and positively correlated with house price growth in 2006 and from 2006 to 2009. The change in search volume in 2006 is also positively and strongly correlated with the growth in the number of mortgage applications for home purchase in 2006. The SVI, in real estate terms, captures very consistently the general interest in the real estate and housing market.

I first show that house price growth and Google SVI are contemporaneously correlated in the year in 2006. I then control for various county-level economic changes including changes in unemployment, changes in mortgage denial rates, changes in mortgage delinquency, growth in income, and growth in house prices during the boom. The OLS estimate of the effect of SVI on house price growth in 2006 barely changes with the inclusion of the controls. Nonetheless, the OLS regression model is subject to two potential endogeneity issues, reverse causality and omitted variable bias. Exogenous shocks in house price growth expectations can drive down real estate investment and consequently interest in the real estate market. Unobservable permanent income shocks can lower demand for housing, which in turn drives house prices

down. Unobservable shocks in the supply of credit can also shift the housing demand curve inward, driving house prices down. Therefore, to identify credibly the effect of lower housing demand due to the increase in mortgage rates on house price growth, one must to address these identification concerns.

I start by estimating an income threshold below which households cannot afford an interest-rate-only mortgage on a average loan in 2005 in the county where they live. I assume that households must have income higher than two times their mortgage payments. I then propose an empirical design that exploits the slope of the income distribution around this affordability threshold to estimate the fraction of households that, after an increase in mortgage interest rates, can no longer afford to pay an interest-rate-only mortgage. Consider an U.S. county in 2005 with a given house price level, average mortgage loan size, and income distribution. Given the average loan size and the interest rate level in 2005, one can estimate the annual payments of an interest-rate-only mortgage and then, given the income distribution, compute the fraction of households that cannot afford to make the mortgage payments. I refer to these households as "mortgaged-out in 2005". Now consider that, between 2005 and 2006, the one-year adjustable rate increased by 1.5%. This increase created an additional fraction of households that could not afford to pay the interest-only 2005 average loan balance. I refer to these households as "mortgaged-out from 2005 to 2006". The proposed empirical design instruments housing demand, measured by the search volume, with the "mortgaged-out from 2005 to 2006", after controlling for "mortgaged-out in 2005". Thus, I exploit the slope of the income distribution around the cutoff of affordability, minimizing dependence on the house price level, which could potentially contaminate the identification. In addition, I control for the growth in house prices between 2000 and 2005.

To test the hypothesis of this chapter, I draw from several data sources. I obtain Google's search volume index from Google Trends. I use two sources of data to measure house prices. The first source is the House Price Index from the Federal Housing Finance Agency, and the second is the Zillow house price index. I use several data sources to measure economic changes at the county level. I use loan data from the Home Mortgage Disclosure Act, income data from the Internal Revenue Service, unemployment data from the Bureau of Labor Statistics, mortgage delinquency data from the Consumer Credit Panel from the Federal Reserve Bank of New York, and income distributions from the American Community Survey. I merge all the datasets at the county level. The final dataset includes approximately 700 counties,

covering approximately 75 million households, containing approximately 180 million people in 2006, around 60% of the U.S. population.

The OLS beta of the effect of changes in 2006 in SVI on contemporaneous growth in house prices without including controls is 0.35 and statistically significant. One standard deviation in interest in the real estate market, as measured by SVI, leads to an increase of 0.35 standards deviations in house prices in 2006. The change in interest in purchasing a home, measured by SVI, has a large economic effect on house price growth. The OLS beta barely changes when the regression includes controls for changes in unemployment, changes in mortgage delinquency, changes in denial rates, growth in income, and growth in house prices during the housing boom. Across the board the OLS beta is fairly stable and varies between 0.35 and 0.40. The IV beta is remarkably close to the OLS beta. This is not surprising considering how stable the OLS coefficient is throughout all the different specifications. The first stage shows that the interest in purchasing a home in 2006 was strongly driven by the change in mortgage interest rates as the first stage shows. The OLS beta is 0.70, with a t-statistic of 8.50.

Although house prices declined in some counties during 2006, it is important to determine whether the 2006 changes did not reversed in the following years. If 2006 house prices growth unwound before the financial crisis, the study of house price behavior in 2006 is irrelevant to understanding the connection between the housing boom and the financial crisis. In this chapter, I show that this is not the case. The change in housing demand is strongly correlated with the house price change from 2006 to 2009. Using the same identification strategy described above, I show that the effect of change in housing demand in 2006 on house price growth from 2006 to 2009 is causal and large. This effect emphasizes the importance of understanding the causes of house price growth in 2006.

This chapter is organized as follows. The remainder of this section reviews related literature. The next section reviews microeconomic indicators, and presents the micro data used in the analysis. The empirical methodology is presented in section III. The results are analyzed in section IV. Section V concludes the chapter.

2.a Related Literature

This chapter aims to fill the gap between two large bodies of literature that have emerged after the financial crisis, the causes of the housing boom and

the causes of the financial crisis. By shedding light in the process through which house prices started a negative trend, it allows others to better model the consequences of the financial crisis. This chapter also relates with works that use the shock in house prices as exogenous to study causal links between other economic outcomes.

The study of the causes of the housing has created a large body of research. Mian and Sufi (2009) and Keys, Mukherjee, Seru, and Vig (2010) seminal works show that in the beginning of the 2000s the U.S. economy experienced an outward shift in the supply of credit. Mian and Sufi (2009) document that less creditworthy borrowers experienced easier access to mortgage credit despite their negative income growth. Keys, Mukherjee, Seru, and Vig (2010) suggest that existing securitization practices adversely affected the screening incentives of subprime lenders. Adelino, Schoar, and Severino (2012) use exogenous changes in the conforming loan limit as an instrument for lower cost of financing and higher supply to show that easier access to credit significantly increases house prices. Agarwal and Ben-David (2014) show that by changing loan officers' compensation structure, from fixed to volume-based pay, more and larger loans are originated, despite leading to higher default rates.

Since this chapter sheds light on house price slow down in 2006, which subsequently led to the great recession, it relates to the literature that studies the consequences of the financial crisis, in particular the large increase in unemployment from 2007 to 2009. Mian and Sufi (2012) claim that the decrease in aggregate demand during the financial crisis was associated with the rapid rise in household debt during the beginning of the 2000s. They argue that the deterioration of households' balance sheets, caused by the decline in house prices, depresses aggregate demand and increases unemployment. Davis, Faberman, and Haltiwanger (2012) argue that the large decrease in aggregate demand led firms to cut back their recruiting intensity, which slowed employment growth. Policy changes, such as the extension of unemployment insurance benefit duration (Rothstein 2011), the rise of government transfers (Mulligan 2012), and the uncertainty of policy direction (Baker, Bloom, and Davis 2012) have also been argued to have negatively affected the recovery of employment.

2.1 Data

In this section, I first discuss some macro indicators that motivate the hypothesis of this chapter. Then, I describe the county-level data, including an extensive description of the Google search data, in particular, how to collect and interpret search data from Google Inc., and how to obtain search data that measures households' interest in the real estate market. Last, I explain how the final dataset used in the remainder of the chapter was constructed.

2.a Macroeconomic Indicators

The top panel in Figure 7 presents the time series of house prices in the U.S.. House prices grew steadily from 1990 until the end of 2005, leveled off during 2006, and started falling sharply after the middle of 2007. The decline in average U.S. house prices from 2007 to 2011 almost reversed the growth from 2000 to 2006.

In general, the number of sales of newly constructed and existing homes followed the same pattern as house prices. However, in 2006, while house prices flattened, the number of homes sales dropped, suggesting an inward shift of housing demand. One striking fact about home transactions is that, during the housing boom, there were over 6 transactions of existing homes for each transaction of a newly constructed home, indicating the importance of the existing housing stock relative to newly constructed homes.

Figure 8 reports the time series for 1-year adjustable-rate and 30-year fixed-rate. From 1999 to 2003 both rates decreased, the 30-year rate by approximately 3 percentage points, and the 1-year rate by approximately 3.5 percentage points. Between the beginning of 2005 and the middle of 2006, the 1-year rate increased by almost 2 percentage points, and the 30-year rate increased by 1 percentage point. If a household had a $300,000 mortgage balance in 2005 and an interest-rate-only mortgage, their annual mortgage payments would have increased by approximately $6,000 in one year. This rapid increase in the 1-year adjustable rate must have affected the ability of households to make their mortgage payments as well as affecting the demand for interest-rate-only and adjustable-rate mortgages; thus it is plausible that the observed decrease in house prices might be related to this increase in interest rates.

Figure 9 reports the time series of the delinquency rate on single-family residential mortgages for the U.S., booked in domestic offices of all com-

mercial banks. During the housing boom, delinquency rates decreased from levels barely over 2 percent to approximately 1 percent in 2004. From the beginning of 2006 through the end of 2006, mortgage delinquency rates only increased by 0.5 percentage points. However, from the end of 2006 until the end of 2009, delinquency rates increased from 2 to 11 percent. Between 2007 and 2009, they leveled off, and after 2012 started to adjust back to pre-crisis levels.

Together, the four figures suggest that the increase in mortgage interest rates between 2005 and 2006 might have shifted the housing demand curve inwards, leading house prices to flatten out during 2006. Other alternative explanations are not excluded, in particular, an exogenous shock in house prices, a credit supply shock, a permanent income shock, or a shock in mortgage delinquency led by the switch from teaser rates to adjustable rates of mortgage vintages from the early 2000s.

2.b Microdata

Google Trends

Internet search engines have become a common and powerful tool for acquiring large quantities of consumer information. Given a string of search terms, a search engine produces a list of matching webpages that are repositories of the desired information. The most popular search engine in the last 10 years is Google, where users can search for virtually anything. The query can be as long as the user desires, but there are gains in efficiency if users use appropriate search terms. Certain search terms are therefore highly representative of the intention of users to acquire information on a specific subject.

Google Inc. uses a sample of servers to track the volume of searches for any term and its geographical origin for every day after January of 2004—provided that there is enough accuracy to estimate the search volume of a given term. This time-series search volume data is formally called the Search Volume Index (SVI).

SVI for a term is available for different countries, states, metropolitan areas, and cities, and reflects the number of searches for that particular term (during a specified period of time and within a region), relative to the total number of searches done at Google (within that region and during that

period). Formally, SVI is defined as:

$$SVI^{j}{}_{r,t} = \frac{SVT^{j}{}_{r,t}}{TSV_t \times MSV_{r,t}} = \frac{SVT^{j}{}_{r,t}}{TSV_t \times \max_{\{q,i\}}\{SVT^{i}{}_{r,q}/TSV_q\}} \quad (1)$$

$SVT^{j}{}_{r,t}$ is the total search volume for term j at period t within region r, TSV_t is the total search volume in Google at time t in region r, and $MSV_{r,t}$ is the maximum of such ratios among the term in the query and within the sample period. Search volume is divided by TSV_t to eliminate any trends that could be present due to a change in the number of Google users, and divided by $MSV_{r,t}$ to scale the time series and not reveal the raw number of searches. Therefore, SVI for a search term is proportional to the percentage of searches for that term during a period of time and within a geographical region.

Google makes the SVI available from its product Google Trends. When querying Google Trends for the search volume of a term, the output is the search volume of all the queries that include that search term. For example, if one requests the search volume of the term "homes", Google Trends will retrieve the search volume of all queries that include the search term "homes", like "homes for sale" or "homes in New York City". For a given term, Google Trends also provides a list of the ten most popular queries that include the search term.

Search terms can be complex—"how to find good real estate deals" and "what are the best real estate websites"—or simple—"remax" and "century 21". Data is available for most simple terms used by people conducting Google searches. Complex terms are less frequently used and usually are not reported in Google Trends.

Another limitation is that Google computes SVI from a random sample of computer servers that varies from day to day and introduces a small sampling error. Since these samples are independent from day to day, I requested data for ten consecutive days and used the average to compute the search volume for a given region. Each time a data request was submitted to Google Trend, the following information was returned: a weekly time series of SVI for the U.S., the SVI for each metropolitan area aggregated for the time period, and the most popular search queries that contained the search terms specified in the request. This studies uses the metropolitan area data.

Housing Demand Proxies I built three different measures of households' interest in purchasing a home, all based on Google search data, but each using different search terms. The first measure uses the specific names of real estate industry firms and websites that are commonly known. A median home buyer looks at 10 homes before purchasing an existing home.[16] To search for houses, either she finds a realtor who researches properties on her behalf, or she screens herself real estate listings. In either case, an internet search engine is highly likely to be used. The list of terms below attempts to capture this part of the purchasing process. The list of search terms includes the names of the largest real estate companies, the largest online real estate listing aggregators, and the words "real estate" and "mls", which are very likely to be present in a general search of real estate listings. The list of search terms for the first housing demand proxy is as follows:

$$\{realestate,\ mls,\ remax,\ century\ 21, realty,\ sothebys,\ era,\ coldwell,\ prudential,\ realtor,\ zillow,\ redfin\}$$

The complete query requested to Google Trends is 'real estate + mls+ remax + century 21 + realty + sothebys + era + coldwell + prudential + realtor + zillow + redfin'. These core terms are present in queries like "remax listings in Los Angeles" or just "remax".

The second proxy for housing demand intends to capture the search volume in the same part of the purchasing process without using company names. This measure alleviates any identification concerns regarding the choice of firm names in the first measure. The second measure then uses the following search terms:

$$\{real\ estate,\ realtor,\ realtors,\ mls,\ realty\}.$$

The complete query requested to Google Trends is: 'real estate + realtor+ realtors + mls + realty'. These core terms are present in queries like "real estate market in Los Angeles" or "mls listings". The third measure exploits general interest in the housing market, and uses the following core terms:

$$\{houses,\ real\ estate,\ construction\}.$$

[16] `http://www.nahb.org/generic.aspx?sectionID=734&genericContentID=196904&channelID=311`

These core terms are present in search queries like "homes for sale" or "new construction in Phoenix". However, because of the generality, they will are also present in queries like "homes for rent" or "rental homes". Thus, I instructed Google Trends to subtract the queries that included "rent" and "rental". As a result, the complete query requested to Google Trends is: 'houses+real estate+construction -rent -rental'. The pluses allow for aggregation and the minuses exclude undesired queries.

Figure 10 depicts the time series of Google search volume for the first measure in four U.S. states: California, Texas, Florida, and Massachusetts. All the time series show the same major pattern. Search volume on real estate terms decreases throughout the financial crisis, and after the crisis levels off to lower level than pre-crisis. Search volume on real estate terms also displays seasonality, typically low at the end of the year, trending up in the beginning of the year, and peaking in the summer. Furthermore, California and Florida have high searches volumes starting in the beginning of the year, during winter time, as opposed to Massachusetts, consistent with the anecdotal evidence that wintertime housing demand is low in places where the winter is harsh. More importantly, the change in search volume from 2004 to 2007 in these four sample states is consistent with the state level response to the financial crisis. That is, search volume decreased faster in states like California and Florida, where the financial crisis was harsher, as opposed to Texas where the crisis was milder.

Figure 11 presents further evidence that search volume of real estate terms is strongly related to housing outcomes. The first two panels show that change in search volume at the county level in 2006 is strongly and positively correlated with house price growth in 2006 and between 2006 and 2009. In areas where search volume for real estate terms dropped during 2006, contemporaneous house prices dropped more intensely, as well as during crisis years. The third panel of Figure 11 shows that the change in search volume in 2006 is positively and strongly correlated with the growth in mortgage applications for home purchase during 2006. Fewer mortgage applications were filed in 2006 when search volume on real estate dropped.

Figures 10 and 11 demonstrate that search volume on real estate terms captures the interest of households in the real estate market very consistently. A similar exercise for the other measures produces the same results. In the remainder of the chapter I use Google search volume as a measure of households' interest in purchasing a home.

Other Data Sources

I use two sources of data to measure house prices. The first source is the House Price Index from the Federal Housing Finance Agency (HPI-FHFA). The HPI-FHFA is a broad measure of the movement of single-family house prices in the U.S.. The HPI-FHFA is a weighted, repeated-sales index, meaning that it measures average price changes in repeated-sales or refinancings on the same properties. I use the Metropolitan Statistical Area House Price Index, which is estimated using sales prices and appraisal data. The HPI-FHFA index is available for nearly all metropolitan areas in the United States.[17]

Loan data is provided by Home Mortgage Disclosure Act (HMDA) provides loan level data from 1990 to 2011. Loan level data is publicly available for lenders that meet a disclosure criteria defined by HMDA every year. Each loan application provides information on *year* of application, *lender*, *type of loan*, *loan amount*, *action taken* by the lender, *reason for denial*, in case the loan is denied, *race*, *sex* and *income* of the applicant and co-applicant, *census tract*, *county FIPS*, and *state FIPS* where the loan was originated, *owner occupancy*, and *purpose*. Loans have four types of purpose: *home purchase*, *home improvement*, *refinancing*, and *multifamily dwelling*. I only use loans that are originated for home purchase and are owner-occupied as principal dwelling.

I use the Internal Revenue Service (IRS) data to compute the county-level average income per capita. The IRS provides county-level data for every year after 1989 until 2012. The provided income data includes adjusted gross income, number of returns, and wage income. In this chapter, income per capita is defined as the ratio of the adjusted gross income to the number of returns.

To measure annual county-level unemployment, I use the Current Population Survey (CPS) by the Bureau of Labor Statistics. Approximately 60,000 households are eligible for the CPS. Sample households are selected by a multistage stratified statistical sampling scheme. A household is interviewed for 4 successive months, then not interviewed for 8 months, then returned to the sample for 4 months after that. I use the official unemployment rate, which is also know as the U3-unemployment rate.

County-level mortgage delinquency and mortgage balances are obtained from the FRBNY website. The data is based on the FRBNY Consumer Credit Panel which constitutes a 5% random sample of the U.S. population

[17]FHFA requires a metro area to have at least 1,000 transactions before it is published.

of individuals who have credit reports with Equifax. Mortgage delinquency is defined as the percent of mortgage debt balance that is over 90 days delinquent. Mortgage debt includes first mortgages, home equity loans and home equity lines of credit. Only counties with an estimated population of at least 10,000 consumers with credit reports in 2010Q4 were included in the dataset.

Finally, I use the American Community Survey (ACS) to obtain the county-level income distribution in 2006. The ACS is an ongoing survey by the U.S. Census that provides data mainly on demographics and income. Approximately, 3 million households are surveyed annually. The ACS data was downloaded from the Social Explorer website.

2.c Construction of the Dataset

Except for the FHFA, HMDA, and Google searches data, all datasets are provided at the county level. The merge of county-level data is based on the county fips and is straightforward. The FHFA uses the Core Based Statistical Area (CBSA) as identifier. I use the NBER's crosswalk to merge CBSA codes to county fips codes. Google Trends provides the full name of the metropolitan area as an identifier. I manually identify the CBSAs that are associated with each metropolitan area from Google, and then proceed to merge the data with using the NBER crosswalk. The number of metros in which there is Google search data available depends on the search terms for which data is required. Given the search terms presented in section 2.b, the final dataset contains between 650 to 700 counties, covering approximately 75 million households, approximately 180 million people in 2006, around 60% of the U.S. population.

2.2 Empirical Methodology

2.a Mechanism

During the beginning of the 2000s while the U.S. economy experienced a housing boom, house price growth reached historical highs, mortgage credit expanded, and many U.S. households with a poor credit history were able to access the mortgage credit markets (Mian and Sufi 2009, Keys et al. 2010, Adelino et al. 2012, Agarwal and Ben-David 2014). Since interest rates stayed at low levels from 2001 until the end of 2004, mortgage products that were sensitive to interest rates, such as interest-rate-only and adjustable-

rate, became popular and some of the most frequently used mortgage products (Mayer, Pence, and Sherlund 2009). Economic growth also followed the housing boom, and in 2004, the Federal Reserve Bank implemented a contractionary monetary policy. Consequently, from the beginning of 2005 until the middle of 2006, the one-year adjustable-rate increased by approximately 1.5%. If a household had a $300,000 mortgage balance in 2005 and an interest-rate-only mortgage, its annual mortgage payments would have gone up by approximately $6,000 in one year. Such a sudden and large increase must have caused great financial stress in households that had mortgage products that were sensitive to interest rates and must have shifted the housing demand curve inwards, particularly, for households that were close to their affordability threshold at current house prices. The first wave of foreclosure fillings in 2006 is an indicator of the inability of these households to bear the increasing interest rates.[18] The hypothesis of this chapter is that the increase in interest rates from 2005 to 2006 lowered the households' interest in purchasing a home, shifting the housing demand curve inward, which then created a downward pressure on house prices from 2005 to 2006.

2.b Measurement

Google Searches on Real Estate Terms It can be challenging to measure demand shocks using equilibrium outcomes, not only because it is difficult to disentangle demand shocks from supply shocks but also because some economic outcomes have in-built inertia—it might take a few months until the sales of homes reflect a shift in the demand curve. Google searches improve these two identification concerns. They are a clear indication of the interest of the user and can be measured almost instantaneously. The interest can, however, sometimes be led by supply factors. For this reason, I design an empirical strategy, presented below, to ensure that Google searches measure a shift in demand through the interest rate channel. To this end, I use Google searches on search terms described in section 2.b to compute a measure of interest in the real estate market. I use three different measures that use different search terms to ensure that the Google searches capture interest in the housing/real estate market. Given the set of terms, as specified in section 2.b, I requested from Google Trends[19] the search volume for 2006,

[18] In 2006 more than 1.2 million foreclosure filling were reported, a 42% increase relative to 2005. http://www.realtytrac.com/press-releases/.

[19] http://www.Google.com/trends/

and the aggregated search volume for 2004 and 2005. Because Google only started collecting search volume in 2004, the data for the first few years is not well populated. Using 2004 and 2005 increases the number of metropolitan areas for which Google has data. Adding the search volume of 2004 and 2005 does not raise an identification concern because 2004 was a housing boom year with low interest rates. I requested data on 10 consecutive days because Google surveys different computer servers each time one requests search volume data in Google Trends, creating a small sampling error. Requesting data on 10 consecutive days eliminates this sampling error. The search volume for a given metropolitan area is, then, the simple average of the 10 data samples. [20] To compute the variation in search volume from 2005 to 2006, I compute the following ratio:

$$\text{SVI}_{\ 05\rightarrow 06,i} = \frac{\text{SVI}_{\ 2006,i}}{\text{SVI}_{\ 2005,i}} - 1,$$

where the search volume index (SVI) is as defined in section 2.b. Intuitively, $\text{SVI}_{\ 2005}$ represents the percentage of all Google searches in 2005 that had terms related to purchasing a house, as specified in section 2.b. The ratio $\text{SVI}_{\ 05\rightarrow 06,i}$ measures the growth in this percentage of searches from 2005 to 2006.

House Price Growth

Using the House Price Index from the Federal Housing Finance Agency, I define the house price growth for county i as follows:

$$\text{House Price Growth}_{\ 2005\rightarrow 2006,i} = \frac{\text{HPI}_{\ 2006,i} - \text{HPI}_{\ 2005,i}}{\text{HPI}_{\ 2005,i}},$$

where HPI is the house price index for the metropolitan area where county i is located.

2.c Estimation Methodology

OLS Estimation

[20] The sampling error disappears after 7 to 8 samples. I did various tests with larger samples to ensure that 10 samples was enough to eliminate the sampling error.

I first show that house price growth and the search volume index (SVI) are contemporaneously correlated in 2006. To this end, I estimate the following regression model for county i:

$$\begin{aligned} \text{House Price}_{\,05\to06,i} &= \beta_0 + \beta_1 \times \text{SVI on Housing}_{\,05\to06,i} \\ &+ \alpha \times X_i + \varepsilon_i, \end{aligned} \tag{2}$$

where X_i is a vector of controls for changes in economic conditions in the county. All variables are standardized, the regression is weighted by the population size of each county, and standard errors are clustered at the state level. To study the stability of the estimated coefficient β_1, I evaluate variations of the model (2) with different subsets of the vector of controls X.

Controls

The vector of controls, X, in (2) contains income per capita growth from 2005 to 2006, *Income per Capita*$_{\{05\to06\}}$, because house prices and searches in Google for housing terms might change due to contemporaneous changes in income. X also includes changes in unemployment from 2005 to 2006, *Unemployment Rate*$_{\{05\to06\}}$, to account for variation that might have been caused by changes in the labor market. Since the changes in house prices might have been caused by omitted variables associated with the exceptional growth in house prices during the housing boom, X also includes a control for the growth in house prices from 2000 to 2005, *FHFA HPI Growth*$_{\{00\to05\}}$. Changes in lending standards might also explain house prices and searches for housing terms, thus the regression model also controls for the change in denial rates for home purchase from 2005 to 2006, *Mortgage Denial Rate*$_{\{05\to06\}}$. Finally, because some early vintages of mortgages might have changed from the teaser rate period to the adjustable-rate period around 2006, the model also controls for the change in mortgage delinquency rate, *Mortgage Delinquency*$_{\{05\to06\}}$.

IV Estimation

The regression model (2) is subject to two potential endogeneity issues, reverse causality and omitted variable bias. Exogenous shocks in house price growth expectations can drive down real estate investment and consequently

interest in the real estate market, as measured by the Google searches. Unobservable permanent income shocks can lower demand for housing, which in turn drives house prices down. Shocks in the supply of credit can also shift the housing demand curve inward, leading house prices down. Therefore, to identify credibly the effect of lower housing demand due to the increase in mortgage rates on house price growth, one needs to address these identification concerns. To this end, I propose an empirical design that exploits the shape of the income distribution to estimate the fraction of households that, after the increase in mortgage interest rates, can no longer afford to pay an interest-rate-only mortgage. I posit that a household cannot afford an interest-rate-only mortgage when the mortgage payments are higher than 50% of the household income.

Consider a U.S. county in 2005 with a given house price level, average mortgage loan size, and income distribution. Given the average loan size and the interest rate level in 2005, one can estimate the annual payments of an interest-rate-only mortgage, and then, given the income distribution, compute the fraction of households that cannot afford to make the mortgage payments. I refer to these households as "mortgaged-out in 2005". Now consider that, between 2005 and 2006, the one-year adjustable-rate increased by 1.5%. This increase created an additional fraction of households that could not afford to pay the interest-only 2005 average loan size. I refer to these households as 'mortgaged-out from 2005 to 2006'. The proposed empirical design instruments housing demand, measured by the search volume, with the "mortgaged-out from 2005 to 2006", after controlling for "mortgaged-out in 2005". Thus, I exploit the slope of the income distribution around the cutoff of affordability, minimizing dependence on the house price level, which could potentially contaminate the identification. In addition, I control for the growth in house prices between 2000 and 2005, and in robustness tests for the price level in 2005.

Figure 13 illustrates this dynamics. Consider that P is the probability distribution of income, r_{2005} is the interest rate level in 2005, $\Delta r > 0$ is the interest rate change from 2005 to 2006, and $\overline{Loan}$ is the average loan in 2005. Then, in Figure 13, the orange area, $P(Income < 2 \times r_{2005} \times \overline{Loan}_{2005})$, represents the fraction of households who cannot afford to pay an interest-rate-only mortgage. When interest rates increase, the new income threshold, $2 \times (r_{2005} + \Delta r) \times \overline{Loan}_{2005}$, is redefined for each county. The fraction of households that are now 'excluded' from the mortgage market, $P(2 \times r_{2005} \times \overline{Loan}_{2005} < Income < 2 \times (r_{2005} + \Delta r) \times \overline{Loan}_{2005})$, the red

area, will depend on the distance between the two income thresholds and on the slope of the income distribution around the old income threshold, $2 \times r_{2005} \times \overline{Loan}_{2005}$. The distance between the income thresholds depends on the size of the interest rate change, which is equal for all counties, and on the initial average loan size. Controlling for the initial size of households "mortgaged-out in 2005", it isolates the dependence of the read area on the slope of the income distribution around the old income threshold. My identification assumption is that the slope of the income distribution around the old income threshold is exogenous to other reasons that drive house price growth from 2005 to 2006.

The first stage works because the higher the fraction of households that cannot afford to pay an interest-only mortgage is, the greater the fall in housing demand is, and, consequently, the greater the fall in number of Google searches is. The exclusion restriction holds assuming that the slope of the income distribution around the affordability threshold is independent of other reasons that drive house prices, particularly exogenous house prices shocks, lending shocks, or permanent income shocks. In the regression model, for county i, I estimate the following in the first-stage:

$$\begin{aligned} \text{SVI on Housing}_{\,2005\to2006,i} &= \theta_0 + \theta_1 \times \text{Mortgaged-out}_{\,2005\to2006,i} \\ &+ \text{Mortgaged-out}_{\,2005,i} + \Theta \times X_i + \eta_i \end{aligned}$$

In the *second stage*, I estimate the following:

$$\begin{aligned} \text{House Price Growth}_{\,2005\to2006,i} &= \beta_0 + \beta_1 \times \widehat{\text{SVI on Housing}}_{\,2005\to2006,i} \\ &+ \text{Mortgaged-out}_{\,2005,i} + \alpha \times X_i + \varepsilon_i. \end{aligned} \tag{3}$$

where X_i is a vector of controls for changes in economic conditions in the county. All variables are standardized, the regression is weighted by the population size of each county, and standard errors are clustered at the state level. To study the stability of the estimated coefficient β_1, I evaluate variations of the model (2) with different subsets of the vector of controls X. Figure 12 depicts a graphed version of the first stage in a scatter by showing that the Google search volume is negatively and strongly correlated with the "mortgaged-out from 2005 to 2006", the fraction of households that were pushed out of the real estate market in 2006 after the increase in mortgage interest rates.

2.3 Results

2.a OLS Results

Table 16 reports the estimation results of model (2) using the first measure of interest in the real estate market, which uses the following search terms:

$$\{real\ estate,\ mls,\ remax,\ century\ 21, realty,\ sothebys,\ era,\ coldwell,\ prudential,\ realtor,\ zillow,\ redfin\}.$$

Column (1) reports the OLS coefficient of house price growth during 2006 on contemporaneous changes in SVI without including controls. The beta is 0.352 with the t-statistic of 2.89. One standard deviation change in interest in the real estate market, as measured by SVI, leads to an increase of 0.35 standard deviations in house prices in 2006. The change in Google searches has a large economic effect on house price growth.

Column (2) reports the coefficient estimates on the control variables when the independent variable is not included. The coefficient on denial rate change from 2005 to 2006 is negative but not statistically significant. The coefficient on change in mortgage delinquency from 2005 to 2006 is negative (-0.258), with a t-statistic of 6.10, implying that higher mortgage delinquency leads to lower house price growth from 2005 to 2006. The coefficient on change in unemployment rate is also negative (-0.145) and statistically significant. Counties that have larger changes in unemployment between 2005 and 2006 have lower contemporaneous house price growth. The coefficient on growth in income per capita is positive and statistically significant, denoting lower house price growth in counties that experience lower income growth. Finally, the growth in house prices from 2000 to 2005 does not statistically correlate with the growth in house prices during 2006. The signs on unemployment change and income growth are consistent with an economic model in which higher permanent income increases demand for housing, which in turn leads to higher house prices.

Column (3) estimates the full model (2). The OLS coefficient on SVI is 0.341; it barely changes relative to the model without controls. After the addition of the controls, the t-statistic increases to 4.28.

Columns (4) and (5) alternate the number of subsets of controls in order to test the stability of the β_1 OLS estimate. In column (4), where only controlling for unemployment, income, and house prices are controlled for,

the estimated coefficient on the searches of real estate terms is 0.364, with a t-statistic of 4.40. In column (5), where only the the mortgage denial rate and the mortgage delinquency are controlled for, the estimated coefficient on SVI is 0.303, with a t-statistic of 2.59.

Finally, columns (6) and (7) show that the estimated coefficient on Google's SVI does not change with the addition "mortgaged-out in 2005". Across the board, the OLS beta is fairly stable and equals approximately 0.35.

2.b IV Results

Since the OLS estimation can suffer from endogeneity issues such as reserve causality or omitted variable bias, I propose an empirical design to address these identification concerns (see in section 2.c). Table 17 presents the results of that empirical design.

Column (1) reports the OLS coefficient of changes in SVI of real estate terms from 2005 to 2006 on the "mortgaged-out from 2005 to 2006". The correlation coefficient is -0.586, with a t-statistic of 8.69, and a R^2 of 34.4%. A one standard deviation change in the "mortgaged-out from 2005 to 2006" leads to 0.586 standard deviations change in the volume of search for real estate terms from 2005 to 2006. The high R^2 shows that "mortgaged-out from 2005 to 2005" explains a large part of the variation of SVI of real estate terms.

Column (2) reports the first stage of the IV estimation. The estimated coefficient changes to -0.708 with a t-statistic of 8.50. The change in the fraction of households that cannot afford an interest-rate-only mortgage, "mortgaged-out from 05 to 06" is a strong instrument for the Google searches on real estate terms. Furthermore, the variation in interest in purchasing a home, as measured by SVI for real estate terms, is strongly driven by the changes in mortgage interest rates.

Column (3) repeats the OLS estimation of column (7) from Table 16 and is only presented for comparison purposes. Column (4) estimates the second stage without any controls, column (5) adds "mortgaged-out in 2005" and the house price growth from 2000 to 2005, column (6) adds to column (5) the controls for unemployment and income growth from 2005 to 2006, and column (7) estimates the full IV model (3).

The estimated coefficients on SVI barely change across specifications of columns (4) to (7). The coefficients vary between 0.383, in the full model, and 0.414 when only "mortgaged-out in 2005" is controlled for. Except for

the house price growth from 2000 to 2005, where the OLS coefficient is biased downwards, the estimated coefficients on the other controls show negligible bias in the OLS estimation.

The IV beta is remarkably close to the OLS beta. This is not surprising considering the stability of the OLS coefficient is throughout all the different specifications. One-standard deviation in the interest in purchasing a home, measured by SVI for real estate terms, leads to a change in house prices of 0.4 standard deviations. The change in interest in purchasing a home in 2006 is strongly driven by the change in mortgage interest rates, as the first stage shows.

2.c Impact on the Financial Crisis

Although house prices declined in some counties during 2006, it is important to determine whether the 2006 changes reversed in the following years. If the 2006 house prices were unwound before the financial crisis, the study of house price behavior in 2006 is irrelevant to understanding the connection between the housing boom and the financial crisis.

Table 22 reports the regression model (2) and (3) with the dependent variable changed to house price growth from 2006 to 2009. Column (1) reports the beta coefficient without controls. The OLS beta on SVI for real estate terms is 0.563 with a t-statistic of 5.93 and a R^2 of 0.317. Alone, the Google searches explain a large amount of the variation of house price growth from 2006 to 2009.

Column (2) only includes the controls. The beta on mortgage denial rate is negative and significant. Like the beta on Google searches, it is also large (-0.534) and very significant: the t-statistic is 11.82. The beta on mortgage delinquency is equal to -0.262 with a t-statistic of 4.95. The beta on unemployment change, although negative, is not statistically different than zero. Income per capita growth is positive and significant.

House price growth during the financial crisis was negative in counties where income grew less. Column (3) reports the full OLS model. The beta on Google searches drops to 0.309 with the inclusion of controls. Column (4) shows that the inclusion of the control "mortgaged-out in 2005" barely changes the OLS estimate. Columns (5) and (6) report the IV estimation. The IV estimate of β_1 is 0.603 with a t-statistic of 2.80. One standard deviation change in Google searches leads to a change of 0.6 standard deviations in house price growth from 2006 to 2009. This is an exceptionally large effect

on house price growth during the financial crisis.

The results in Table 22 indicate that the effect on house price growth in 2006 not only unwound during the crisis, but in fact it amplified, making the study of early house price behavior more relevant to understanding the outcomes during the financial crisis.

2.d Robustness Tests

Tables 18 to 19 present the OLS and IV results for the second proxy of housing demand, respectively. The SVI for this proxy uses the following search terms:

$$\{real\ estate,\ realtor,\ realtors,\ mls,\ realty\}.$$

The OLS beta of the effect of SVI on house price growth equals 0.384 with a t-statistic of 3.19, without controls. The inclusion of all controls changes the OLS beta to 0.415 with a t-statistic of 5.38. Columns (4) to (7) report the estimation of model (2) with different subsets of X. Across the board, the OLS beta is fairly stable and equals approximately 0.40.

Column (1) of Table 19 reports the OLS coefficient of changes in SVI of real estate terms from 2005 to 2006 on the "mortgaged-out from 2005 to 2006". The correlation coefficient is -0.644, with a t-statistic of 6.17, and a R^2 of 41.4%. Column (2) reports the first stage of the IV estimation. The estimated coefficient changes to -0.801 with a t-statistic of 10.63. The estimated coefficients on SVI barely change across specifications of columns (4) to (7). The coefficients vary between 0.343, in column (6), and 0.377 when only "mortgaged-out in 2005" and house price growth from 2000 to 2005 are controlled for. The IV beta is slightly smaller than the OLS beta. One-standard deviation in the interest in purchasing a home, measured by SVI for real estate terms, leads to a change in house prices of 0.373 standard deviations. Although the terms in this second proxy are more general than the firm names of the first proxy, all the results still hold.

Tables 20 to 21 present the results for the third proxy of housing demand. The SVI for this proxy uses the following search terms:

$$\{houses,\ real\ estate,\ construction\}.$$

The OLS beta of the effect of SVI on house price growth equals 0.409 with a t-statistic of 4.34, without controls. The inclusion of all controls changes the

OLS beta to 0.446 with a t-statistic of 5.91. Columns (4) to (7) report the estimation of model (2) with different subsets of X. Across the board, the OLS beta is fairly stable and equals approximately 0.42. Column (1) of Table 21 reports the OLS coefficient of changes in SVI of real estate terms from 2005 to 2006 on the "mortgaged-out from 2005 to 2006". The correlation coefficient is -0.672, with a t-statistic of 8.85, and a R^2 of 45.1%. Column (2) reports the first stage of the IV estimation. The estimated coefficient changes to -0.623 with a t-statistic of 7.61. The IV beta is slightly larger than the OLS beta. One-standard deviation in the interest in purchasing a home, measured by SVI for real estate terms, leads to a change in house prices of 0.46 standard deviations. Although this proxy is even more general than the previous two, the results still hold.

Table 23 presents a robustness test, where housing growth from 2000 to 2005 is replaced by the level of house prices in 2005. The baseline methodology controls for "mortgaged-out in 2005" to account for the variation in the "mortgaged-out from 2005 to 2006" that can be related to the level of house prices in 2005. This robustness test assesses if there is any variation in the price level that was not account for in the baseline specification. Table 23 shows that the results of do not change if the level of house prices is controlled for. The estimated IV beta equals 0.454 with a t-statistic of 2.42. Table 23 also reports the first stage of the IV estimation and the OLS results. All columns hold similar results than the previous specifications.

2.4 Concluding Remarks

The extraordinary growth in house prices in the beginning of the 2000s lasted until the end of 2005. During 2006, the U.S. average house price leveled off. However, across cities, house price growth in 2006 was remarkably different. For instance, in Baltimore City, MD, house prices grew by 19.2% during 2006, while in Santa Barbara County, CA, they dropped by 11.3%. What caused this heterogeneity in house prices across cities in 2006? Was it related with the contractionary monetary policy that started in the end of 2004?

This chapter studies the effect of interest rate changes on housing demand at the end of a housing boom and the subsequent effect on house prices. I build three different proxies for housing demand using Google search volume on search terms such as "remax", "construction", or "real estate", which

are likely to be searched for during the process of purchasing a home. One-standard deviation increase on the change in interest in purchasing a home form 2005 to 2006, measured by the Google search volume, leads to a 0.4 standard deviation increase in contemporaneous house prices growth.

To causally identify this effect and the channel, I first compute a household income threshold for each county defined by the necessary income to afford an interest-rate-only mortgage on an average loan in 2005. I then exploit the slope of the county's income distribution around this affordability threshold to estimate the fraction of households that, after and increase in mortgage interest rates from 2005 to 2006, can no longer afford to pay an interest-rate-only mortgage. I use this fraction of households as an instrumental variable for housing demand.

The change in interest in purchasing a home in 2006 is strongly driven by the change in mortgage interest rates, as the first stage shows. The IV beta is remarkably close to the OLS beta. This chapter sheds light on the transition process between the housing boom in the 2000s and the subsequent financial crisis, and the impact of monetary policy on housing demand and house prices at the end of a housing boom.

References

Adelino, M., A. Schoar, and F. Severino (2012, February). Credit supply and house prices: Evidence from mortgage market segmentation. Working Paper 17832, National Bureau of Economic Research.

Adsera, A. (2005). Vanishing children: From high unemployment to low fertility in developed countries. *The American Economic Review 95*(2), pp. 189–193.

Agarwal, S. and I. Ben-David (2014, February). Do loan officers incentives

lead to lax lending standards? Working Paper 19945, National Bureau of Economic Research.

Angrist, J., V. Lavy, and A. Schlosser (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics 28*(4), pp. 773–824.

Angrist, J. D. and W. N. Evans (1998). Children and their parents' labor supply: Evidence from exogenous variation in family size. *The American Economic Review 88*(3), pp. 450–477.

Angrist, J. D. and A. B. Krueger (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *The Journal of Economic Perspectives 15*(4), 69–85.

Baker, S. R., N. Bloom, and S. J. Davis (2012). Has Economic Policy Uncertainty Hampered the Recovery? Working Papers 2012-003, Becker Friedman Institute for Research In Economics.

Barro, R. J. and G. S. Becker (1989). Fertility choice in a model of economic growth. *Econometrica 57*(2), pp. 481–501.

Becker, G. S. (1960). An economic analysis of fertility. In *Demographic and Economic Change in Developed Countries*, pp. 209–240. Columbia University Press.

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal 75*(299), pp. 493–517.

Becker, G. S. and H. G. Lewis (1973). On the interaction between the quantity and quality of children. *The Journal of Political Economy 81*(2), S279–S288.

Becker, G. S., K. M. Murphy, and R. Tamura (1990). Human capital, fertility, and economic growth. *Journal of Political Economy 98*(5), S12–37.

Bertocchi, G., M. Brunetti, and C. Torricelli (2011). Marriage and other risky assets: A portfolio approach. *Journal of Banking & Finance 35*(11), 2902–2915.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005). The more the merrier? the effect of family size and birth order on children's education. *The Quarterly Journal of Economics 120*(2), 669–700.

Butz, W. P. and M. P. Ward (1979). The emergence of countercyclical u.s. fertility. *The American Economic Review 69*(3), pp. 318–328.

Charles, K. K., E. Hurst, and M. J. Notowidigdo (2013, April). Manufacturing decline, housing booms, and non-employment. (18949).

Choi, H. and H. Varian (2012). Predicting the present with google trends. *Economic Record 88*, 2–9.

Da, Z., J. Engelberg, and P. Gao (2011). In search of attention. *The Journal of Finance 66*(5), 1461–1499.

Davis, S. J., R. J. Faberman, and J. C. Haltiwanger (2012, January). Recruiting Intensity during and after the Great Recession: National and Industry Evidence. Working Paper 17782, National Bureau of Economic Research.

Dettling, L. J. and M. S. Kearney (2011, October). House prices and birth rates: The impact of the real estate market on the decision to have a baby. Working Paper 17485, National Bureau of Economic Research.

Ermisch, J. F. (1989). Purchased child care, optimal family size and mother's employment: Theory and econometric analysis. *Journal of Population Economics 2*(2), pp. 79–102.

Felson, M. and M. Solaun (1975). The fertility-inhibiting effect of crowded apartment living in a tight housing market. *American Journal of Sociology 80*(6), pp. 1410–1427.

Ferreira, F. and J. Gyourko (2011, August). Anatomy of the beginning of the housing boom: U.s. neighborhoods and metropolitan areas, 1993-2009. Working Paper 17374, National Bureau of Economic Research.

Galor, O. and D. N. Weil (1996). The gender gap, fertility, and growth. *The American Economic Review 86*(3), pp. 374–387.

Galor, O. and D. N. Weil (2000). Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *American economic review*, 806–828.

Gertler, P., D. I. Levine, and E. Moretti (2009). Do microfinance programs help families insure consumption against illness? *Health Economics 18*(3), 257–273.

Glaeser, E. L., J. D. Gottlieb, and K. Tobio (2012). Housing booms and city centers. *American Economic Review 102*(3), 127–33.

Goel, S., J. M. Hofman, S. Lahaie, D. M. Pennock, and D. J. Watts (2010). Predicting consumer behavior with web search. *Proceedings of the National Academy of Sciences 107*(41), 17486–17490.

Greenwood, R. and D. Scharfstein (2013). The growth of finance. *Journal of Economic Perspectives 27*(2), 3–28.

Guerrieri, V., D. Hartley, and E. Hurst (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics 100*(0), 45 – 60.

Jappelli, T. and L. Pistaferri (2011). Financial integration and consumption smoothing. *The Economic Journal 121*(553), 678–706.

Joseph Hotz, V., J. A. Klerman, and R. J. Willis (1997). The economics of fertility in developed countries. *Handbook of population and family economics 1*, 275–347.

Kaplan, S. N. and J. Rauh (2010). Wall street and main street: What contributes to the rise in the highest incomes? *Review of Financial Studies 23*(3), 1004–1050.

Karlan, D. and J. Zinman (2010). Expanding credit access: Using randomized supply decisions to estimate the impacts. *Review of Financial Studies 23*(1), 433–464.

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

Kohler, H.-P., J. R. Behrman, and A. Skytthe (2005). Partner + children = happiness? the effects of partnerships and fertility on well-being. *Population and Development Review 31*(3), 407–445.

Kulu, H. and A. Vikat (2007). Fertility differences by housing type: The effect of housing conditions or of selective moves? *Demographic Research S6*(26), 775–802.

Leventhal, T. and S. Newman (2010). Housing and child development. *Children and Youth Services Review 32*(9), 1165 – 1174.

Levine, R. and Y. Rubinstein (2013, August). Liberty for more: Finance and educational opportunities. Working Paper 19380, National Bureau of Economic Research.

Love, D. (2009). The effects of marital status and children on savings and portfolio choice. *Review of Financial Studies 23*(1), 385–432.

Malthus, T. R. (1798). *An essay on the principle of population*, Volume 1. Norton, New York.

Margolis, R. and M. Myrskyla (2011). A global perspective on happiness and fertility. *Population and Development Review 37*(1), 29–56.

Mayer, C., K. Pence, and S. M. Sherlund (2009). The rise in mortgage defaults. *Journal of Economic Perspectives 23*(1), 27–50.

Melzer, B. T. (2011). The real costs of credit access: Evidence from the payday lending market. *The Quarterly Journal of Economics 126*(1), 517–555.

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

Mian, A. R. and A. Sufi (2012, February). What explains high unemployment? the aggregate demand channel. (17830).

Mincer, J. (1963). Market prices, opportunity costs, and income effects. *Measurement in economics*, 67–82.

Morse, A. (2011). Payday lenders: Heroes or villains? *Journal of Financial Economics 102*(1), 28–44.

Mulder, C. H. and F. C. Billari (2010). Homeownership regimes and low fertility. *Housing Studies 25*(4), 527–541.

Mulligan, C. B. (2012). *The Redistribution Recession: How Labor Market Distortions Contracted the Economy*. Oxford University Press.

Parrado, E. and S. Morgan (2008). Intergenerational fertility among hispanic women: New evidence of immigrant assimilation. *Demography 45*(3), 651–671.

Philippon, T. and A. Reshef (2012). Wages and human capital in the u.s. finance industry: 1909-2006. *The Quarterly Journal of Economics*.

Philippon, T. and A. Reshef (2013). An international look at the growth of modern finance. *Journal of Economic Perspectives 27*(2), 73–96.

Rothstein, J. (2011, October). Unemployment Insurance and Job Search in the Great Recession. Working Paper 17534, National Bureau of Economic Research.

Schultz, T. P. (1985). Changing world prices, women's wages, and the fertility transition: Sweden, 1860-1910. *Journal of Political Economy 93*(6), pp. 1126–1154.

Shu, P. (2013). Career choice and skill development of mit graduates: Are the best and brightest going into finance?

Staiger, D. and J. H. Stock (1997, May). Instrumental variables regression with weak instruments. *Econometrica 65*(3), 557–586.

Strom, S. (2010). Housing and first births in sweden, 1972-2005. *Housing Studies 25*(4), 509–526.

Thomson, E. and Y. Brandreth (1995). Measuring fertility demand. *Demography 32*(1), 81–96.

Vlastakis, N. and R. N. Markellos (2012). Information demand and stock market volatility. *Journal of Banking & Finance 36*(6), 1808–1821.

Willis, R. J. (1973). A new approach to the economic theory of fertility behavior. *Journal of Political Economy 81*(2), pp. S14–S64.

Figure 1: Fertility Rate and Mortgage Origination in US from 1990 to 2010

The top panel shows the aggregate number of births (in millions) in the U.S. between 1990 and 2000. The mid panel presents the aggregate fertility rate for women in the U.S. in the same period. Fertility rate is defined as the total number of births divided by the number of women (in thousands) with ages between 15 and 44. The aggregate data on births is from the Centers for Disease Control and Prevention (`http://www.cdc.gov/nchs/births.htm`). The bottom panel shows the aggregate value in millions of dollars of originations for home purchase from HMDA between 1990 to 2010.

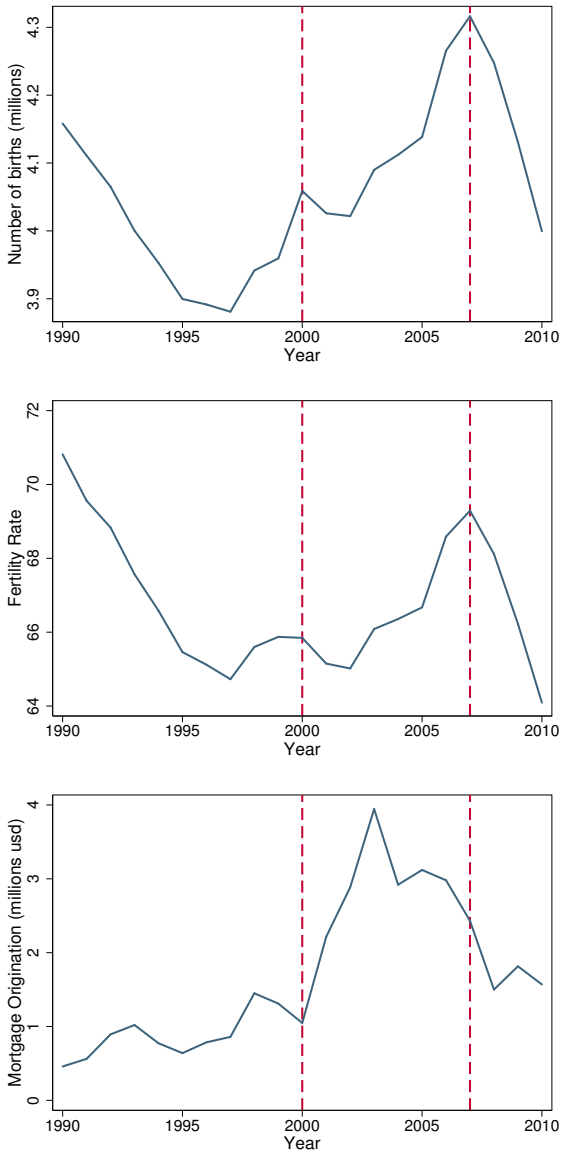

Figure 2: Annual Sales of Newly Constructed and Existing Homes in US

The top panel depicts the annual number of house transactions for newly constructed and the bottom panel reports the annual number of house transactions for existing homes. Both levels are presented in millions. Data was collected from the National Association of Realtors website.

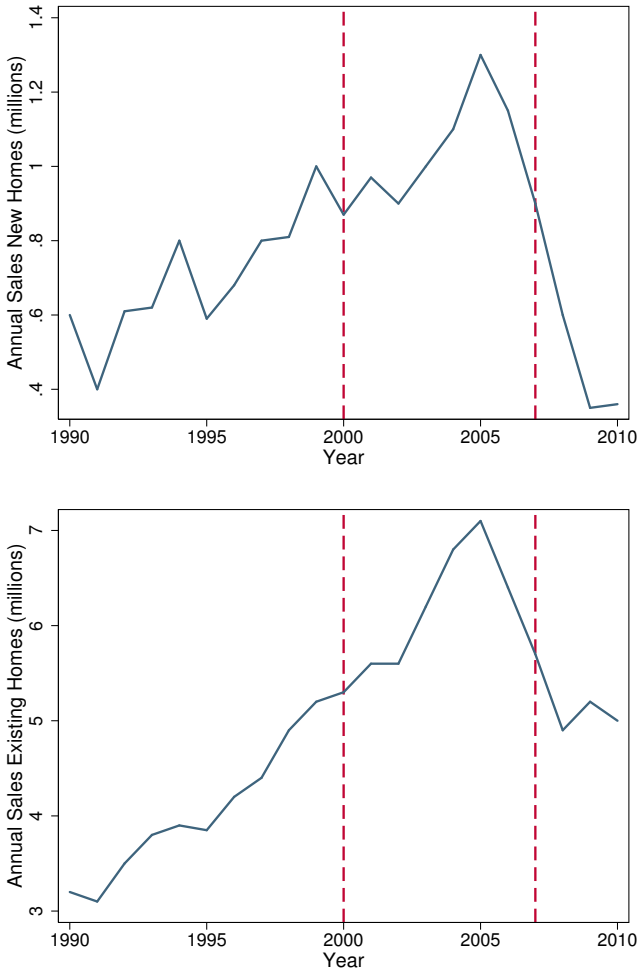

Figure 3: Fertility and Mortgage Origination in the Cross-Section: County-Level

The picture below shows the correlation between the fertility change and the per capita mortgage origination change. Fertility rate is defined as the number of births divided by the number of women (in thousands) of ages between 15 and 44. Per capita mortgage origination change is defined as the mortgage origination for home purchase divided by the population (in thousands) in the county. Each dot represents a county and the size is proportional to the size of the population in the county. Only counties with more than 100,000 people are plotted in the graph. The red line is the fitted regression line. The left panel depicts the relationship between mortgage origination and fertility rates from 1995 to 2000, the middle panel from 2000 to 2006 and the right panel from 2007 and 2010.

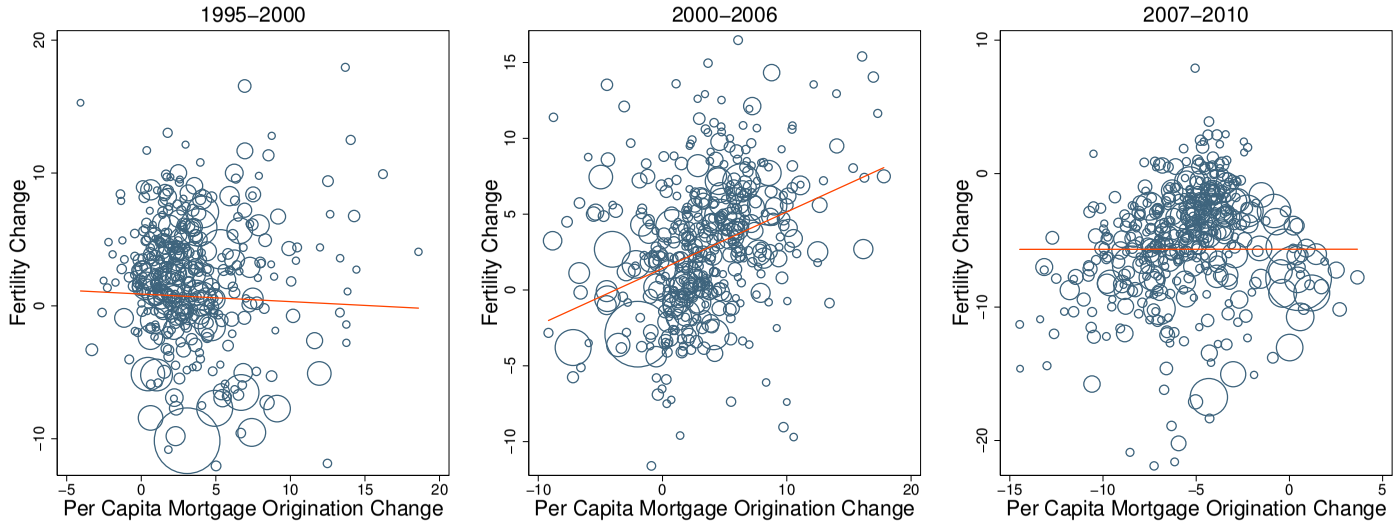

Figure 4: Fertility Rate in Top and Bottom Quintile of Credit Growth

This figure presents the fertility rates for two groups of counties. Counties were sorted by the change in per capita mortgage origination for home purchase. Mortgage credit origination comes from HMDA and is defined as the total number of mortgage loans originated for home purchase. The blue and green lines are the averages for the top and bottom quintiles, respectively. The red dashed lines are the bounds for the 95% confidence intervals. Fertility rate is defined as the total number of births divided by the number of women (in thousands) with ages between 15 and 44.

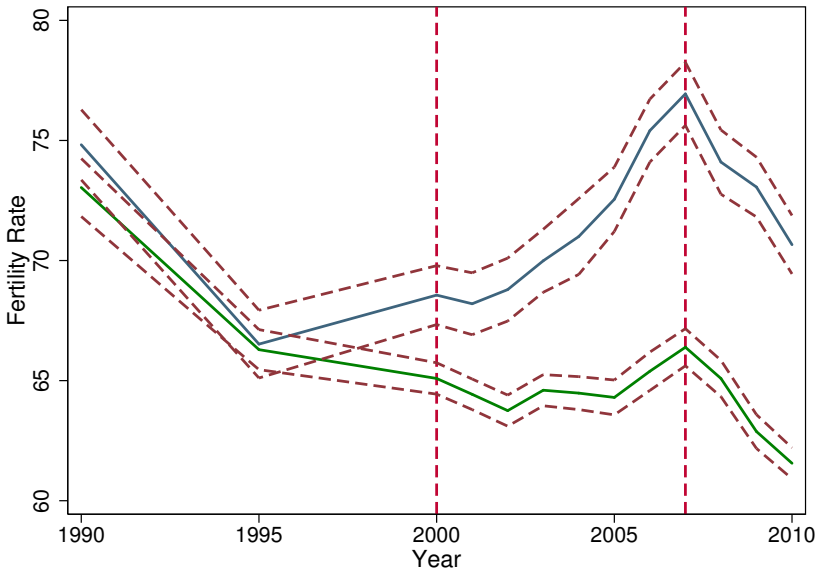

Figure 5: Rent Price by Apartment Size

The picture below shows the asked price per bedroom by apartment size. Apartment size is measured by the number of bedrooms. Asked prices are estimated from 30 million craigslist adds from 2008 to 2013 for all the cities where craigslist is present in the US.

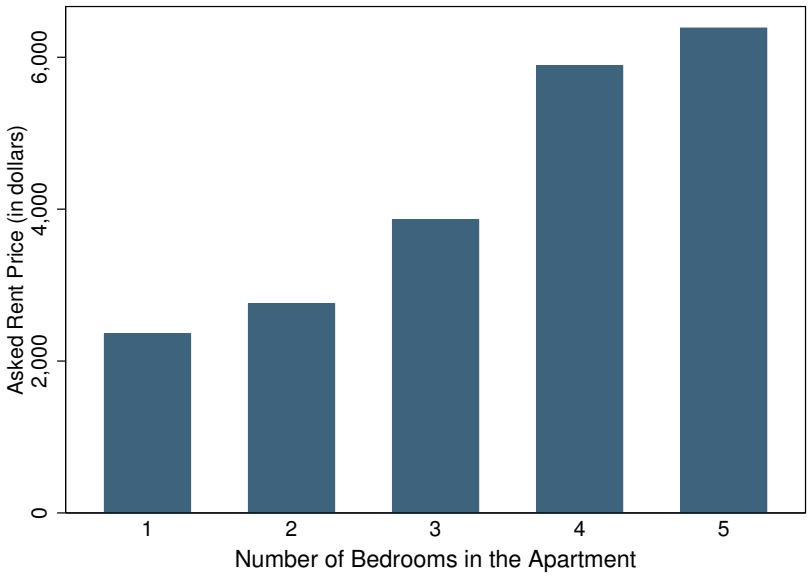

Figure 6: Changes in Homeownership Rate by Age Group during the Housing Boom

In the picture below each bar represents the change in percentage points of homeownership rate by age group. The blue bars report the changes from 2000 to 2005 and the light yellow the changes from 1990-1995. Homeownership rate is defined as the proportion of owner households to the total number of occupied households. The estimates are based on the Current Population Survey and Housing Vacancy Survey.

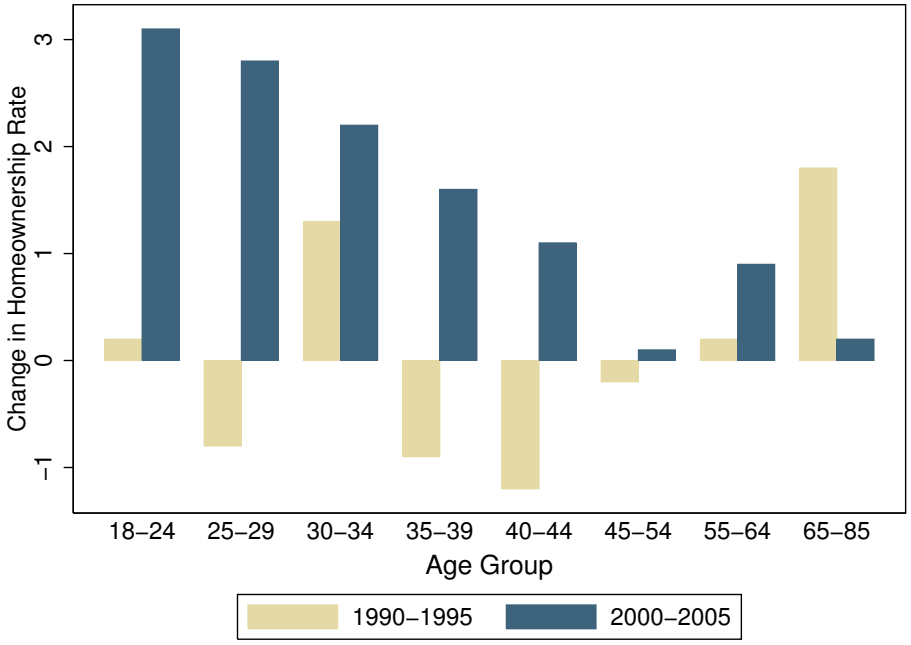

Table 1: Summary of Statistics: County Level

Per Cap Origination is defined as the mortgage origination for home purchase divided by the population (in thousands) in the county. Fertility rate is defined as the number of births divided by the number of women (in thousands) with ages between 15 and 44. IRS Income growth is computed using the adjust gross income from the IRS data. Female College is the fraction of women in the county who have at least a bachelor's degree. Unemployment is from the BLS. Homeownership is from the Census. Construction is the number of new building permits for single homes from the Census. HP is the home price index from the FHFA.

	N	Mean	Std	10th	50th	90th
Per Cap Origination$_{\{00\to06\}}$	700	4.35	3.87	-0.44	4.20	8.99
HP$_{\{00\to06\}}$ × Ownership$_{\{00\}}$	700	0.38	0.26	0.14	0.25	0.75
HP growth$_{\{00\to06\}}$	700	0.57	0.39	0.21	0.39	1.14
Homeownership$_{\{2000\}}$	700	0.67	0.079	0.55	0.66	0.77
House Units Growth$_{\{00\to06\}}$	696	0.10	0.50	-0.48	0.069	0.69
Per Capita IRS Inc. Growth$_{\{00\to06\}}$	700	0.082	0.069	0.010	0.088	0.16
Per Capita IRS Inc.$_{\{2000\}}$	700	49.8	13.7	37.6	47.1	62.9
Unemployment$_{\{00\to06\}}$	700	0.83	1.08	-0.40	0.90	2.20
Female College$_{\{00\to06\}}$	700	0.043	0.015	0.025	0.043	0.060
Fertility$_{\{00\to06\}}$	700	3.40	4.40	-2.10	3.35	8.20
Fertility Change $_{\{95\to00\}}$	700	2.03	5.56	-4.92	2.35	8.01
Fertility Rate $_{\{2000\}}$	700	66.6	10.6	54.7	66.2	78.0

Table 2: Summary of Statistics: Zip Code Level

Per Cap Origination is defined as the mortgage origination for home purchase divided by the population (in thousands) in the zip code. Fertility rate is defined as the number of births divided by the number of women (in thousands) with ages between 15 and 44. IRS Income growth is computed using the adjust gross income from the IRS data. Income and Capital Gains are reported in thousands of dollars. HP is house price growth from Zillow. Female College is the fraction of women in the zip code who have at least a bachelor's degree. Demographics and Unemployment are from the Census and the American Community Survey (ACS). Employment is computed using the data from the County of Business Patterns. House units growth is computed using the number of house units per zip code as reported by the Census and ACS. Old-Homeowners is the instrumental variable that measures fraction of homeowners who are older than 65 years-old and live alone, while 75-Homeowners measures the fraction of homeowners who are older than 75 years-old and live alone, both described and discussed in section 1.c.

	N	Mean	Std	10th	50th	90th
Homeownership$_{\{2000\}}$	2753	0.65	0.19	0.36	0.69	0.86
HP Growth$_{\{00\rightarrow06\}}$	2753	1.06	0.54	0.35	1.06	1.78
House Units Growth$_{\{00\rightarrow06\}}$	2753	0.13	0.23	-0.014	0.066	0.31
Per Cap Origination$_{\{00\rightarrow06\}}$	2753	4.70	12.0	-4.96	3.10	15.9
Female Unemployment$_{\{00\rightarrow06\}}$	2753	0.011	0.036	-0.026	0.014	0.047
Female College$_{25\text{-}44,\{00\rightarrow06\}}$	2753	-0.017	0.046	-0.075	-0.015	0.035
Employment Change$_{\{00\rightarrow06\}}$	2753	0.0021	0.19	-0.076	0.0088	0.089
Per Capita Inc. Growth$_{\{00\rightarrow06\}}$	2753	0.22	0.16	0.076	0.19	0.40
Per Capita Income$_{\{2000\}}$	2753	52.72	41.04	28.00	42.65	83.42
Capital Gains$_{\{2006\}}$	2753	7.96	22.46	0.53	2.56	16.26
Fraction of Hisp.$_{\{00\rightarrow06\}}$	2753	0.028	0.041	-0.0084	0.020	0.083
Fraction of Black.$_{\{00\rightarrow06\}}$	2753	0.0033	0.033	-0.022	0.0021	0.034
Population$_{\{2000\}}$	2753	24,793	18,142	4,951	21,367	48,227
Population$_{\{2006\}}$	2753	26,732	19,158	5,239	23,673	51,985
Fertility$_{\{2000\}}$	2753	62.1	17.9	41.5	60.4	84.7
Fertility$_{\{2006\}}$	2753	65.5	21.9	40.6	63.6	92.5
Fertility$_{\{00\rightarrow06\}}$	2753	3.37	14.1	-11.6	2.40	19.0
Old-Homeowners$_{\{2000\}}$	2753	.099	.052	0.041	0.092	0.158
75-Homeowners$_{\{2000\}}$	2753	.057	.038	0.020	0.052	0.98

Table 3: Homeownership Change and the Old-Homeowners Instrument: Zip Code Level

This table presents the regression coefficients of the instrument defined by the fraction of homeowners living alone and older than 65 years-old, as explained in section 1.c, with the change in homeownership for different age groups, in panel A, with other shocks and levels, in panel B and C. Per Cap Origination is defined as the mortgage origination for home purchase divided by the population (in thousands) in the zip code. Fertility rate is defined as the number of births divided by the number of women (in thousands) with ages between 15 and 44. IRS Income growth is computed using the adjust gross income from the IRS data. HP is house price growth from Zillow. Female College is the fraction of women in the zip code who have at least a bachelor's degree. Demographics and Unemployment are from the Census and the American Community Survey (ACS). Employment is computed using the data from the County of Business Patterns. Bedrooms growth is computed using the number of bedrooms per zip code as reported by the Census and ACS. All regressions have county effects. Standard errors are robust and clustered at the state level. ***,**,* coefficient estimate statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; t-statistics in parentheses.

Panel A

	Ownership Change from 2000 to 2006						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	25< age <34	35< age <44	45< age <54	55< age <59	60< age <64	65< age <74	age >75
Old-Homeowners$_{\{2000\}}$	0.34***	0.46***	0.06	0.05*	0.00	-0.35***	-0.49***
	(4.81)	(9.43)	(0.69)	(2.30)	(0.09)	(-6.47)	(-11.07)
County Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Zip Codes	2753	2753	2753	2753	2753	2753	2753
R-squared	0.218	0.232	0.144	0.122	0.123	0.237	0.319

Panel B

	From 1996 to 2000	From 2000 to 2006			Levels in 2000	
	(1)	(2)	(3)	(4)	(5)	(6)
	HP Growth	Bedrooms Growth	HP Growth	Inc. pc Growth	Log HP	Log pc Income
Old-Homeowners$_{\{2000\}}$	-0.05	-0.17***	0.11**	0.20***	-0.20**	-0.13**
	(-0.76)	(-4.15)	(2.60)	(7.49)	(-3.26)	(-2.46)
County Effects	Yes	Yes	Yes	Yes	Yes	Yes
#Zip Codes	2676	2753	2753	2753	2753	2753
R-squared	0.034	0.030	0.269	0.056	0.094	0.018

Panel C

	From 2000 to 2006					
	(1)	(2)	(3)	(4)	(5)	(6)
	Fraction Hisp.	Fraction Black	Male Unemp.	Female Unemp.	Emp. Change	Female College
Old-Homeowners$_{\{2000\}}$	-0.04	-0.00	0.01	0.02	-0.01	0.26***
	(-1.28)	(-0.07)	(0.23)	(0.49)	(-0.27)	(6.60)
County Effects	Yes	Yes	Yes	Yes	Yes	Yes
#Zip Codes	2753	2753	2753	2753	2753	2753
R-squared	0.019	0.001	0.010	0.005	0.002	0.066

Table 4: Homeownership Change and the 75-Homeowners Instrument: Zip Code Level

This table presents the regression coefficients of the instrument defined by the fraction of homeowners living alone and older than 75 years-old, as explained in section 1.c, with the change in homeownership for different age groups, in panel A, with other shocks and levels, in panel B and C. Per Cap Origination is defined as the mortgage origination for home purchase divided by the population (in thousands) in the zip code. Fertility rate is defined as the number of births divided by the number of women (in thousands) with ages between 15 and 44. IRS Income growth is computed using the adjust gross income from the IRS data. HP is house price growth from Zillow. Female College is the fraction of women in the zip code who have at least a bachelor's degree. Demographics and Unemployment are from the Census and the American Community Survey (ACS). Employment is computed using the data from the County of Business Patterns. Bedrooms growth is computed using the number of bedrooms per zip code as reported by the Census and ACS. All regressions have county effects. Standard errors are robust and clustered at the state level. ***,**,* coefficient estimate statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; t-statistics in parentheses.

Panel A

	Ownership Change from 2000 to 2006						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	25< age <34	35< age <44	45< age <54	55< age <59	60< age <64	65< age <74	age >75
75-Homeowners$_{\{2000\}}$	0.31***	0.40***	0.09	0.07**	0.02	-0.22***	-0.54***
	(4.22)	(5.21)	(0.95)	(2.94)	(0.79)	(-4.17)	(-17.84)
County Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Zip Codes	2753	2753	2753	2753	2753	2753	2753
R-squared	0.208	0.216	0.146	0.124	0.123	0.195	0.373

Panel B

	From 1996 to 2000	From 2000 to 2006			Levels in 2000	
	(1)	(2)	(3)	(4)	(5)	(6)
	HP Growth	Bedrooms Growth	HP Growth	Inc. pc Growth	Log HP	Log pc Income
75-Homeowners$_{\{2000\}}$	-0.03	-0.17***	0.08	0.18***	-0.15**	-0.09*
	(-0.51)	(-4.02)	(1.84)	(6.13)	(-2.60)	(-1.98)
County Effects	Yes	Yes	Yes	Yes	Yes	Yes
#Zip Codes	2676	2753	2753	2753	2753	2753
R-squared	0.033	0.030	0.263	0.049	0.077	0.011

Panel C

	From 2000 to 2006					
	(1)	(2)	(3)	(4)	(5)	(6)
	Fraction Hisp.	Fraction Black	Men Unemp.	Women Unemp.	Emp. Change	Women College
75-Homeowners$_{\{2000\}}$	-0.03	0.01	0.02	0.02	0.00	0.23***
	(-1.03)	(0.22)	(0.41)	(0.36)	(0.06)	(7.72)
County Effects	Yes	Yes	Yes	Yes	Yes	Yes
#Zip Codes	2753	2753	2753	2753	2753	2753
R-squared	0.018	0.001	0.010	0.005	0.002	0.053

Table 5: Mortgage Origination and Fertility During the Housing Boom: County Level

Per Cap Origination is defined as the mortgage origination for home purchase divided by the population (in thousands) in the zip code. Fertility rate is defined as the number of births divided by the number of women (in thousands) with ages between 15 and 44. IRS Income growth is computed using the adjust gross income from the IRS data. The bottom row indicates which regressions have state effects and population weights. All regressions include controls for age and demographic in levels and differences as described in section 1.c. HP is house price growth from FHFA. Construction Growth is the growth in building permits from the census records. Female college is the fraction of females in the county who have at least a bachelor's degree. Standard errors are robust and clustered at the MSA level. ***,**,* coefficient estimate statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; standard errors in parentheses.

	Change in Fertility from 2000 to 2006									
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) OLS	(10) OLS
Per Cap Origination$_{\{00\to06\}}$		0.46***	0.41***		0.37***	0.35***	0.37***	0.30***		0.30***
		(0.07)	(0.07)		(0.07)	(0.06)	(0.06)	(0.05)		(0.06)
HP$_{\{00\to06\}}$ × Ownership$_{\{00\}}$									20.97***	22.09***
									(7.29)	(7.27)
HP growth$_{\{00\to06\}}$									-15.53***	-16.63***
									(5.27)	(5.21)
Homeownership$_{\{2000\}}$									-22.68***	-20.15***
									(5.62)	(5.78)
House Units Growth$_{\{00\to06\}}$									0.94**	-0.11
									(0.39)	(0.41)
Per Capita Inc. Growth$_{\{00\to06\}}$	3.18		-2.15	-2.08		-5.92**		-7.37**	-2.62	-4.90
	(3.12)		(3.08)	(2.79)		(2.84)		(3.08)	(3.16)	(3.08)
Per Capita Inc.$_{\{2000\}}$	-0.03		-0.01	-0.07***		-0.04*		-0.01	-0.03*	-0.02
	(0.03)		(0.03)	(0.02)		(0.03)		(0.02)	(0.02)	(0.02)
Unemployment$_{\{00\to06\}}$	-0.63**		-0.50**	-1.53***		-1.49***		-1.29***	-1.48***	-1.43***
	(0.25)		(0.25)	(0.38)		(0.39)		(0.32)	(0.34)	(0.33)
Female College$_{\{00\to06\}}$	4.07		1.83	4.99		0.35		-4.01	2.48	-0.87
	(10.86)		(10.44)	(9.31)		(9.39)		(10.36)	(10.26)	(10.10)
Fertility Change $_{\{95\to00\}}$	-0.06		-0.04	-0.06		-0.04		-0.06	-0.06	-0.06
	(0.05)		(0.05)	(0.04)		(0.04)		(0.04)	(0.05)	(0.05)
Fertility Rate $_{\{2000\}}$	-0.02		-0.06	-0.13***		-0.14***		-0.08***	-0.08**	-0.08***
	(0.04)		(0.04)	(0.04)		(0.04)		(0.03)	(0.03)	(0.03)
State Effects	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop Weight	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Age-Demo Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Counties	700	700	700	700	700	700	700	700	696	696
R-squared	0.137	0.097	0.199	0.369	0.290	0.403	0.344	0.505	0.502	0.525

Table 6: Mortgage Origination and Fertility During The Boom: Zip Code Level - Old-Homeowners

Per Cap Origination is defined as the mortgage origination for home purchase divided by the population (in thousands) in the zip code. Fertility rate is defined as the number of births divided by the number of women (in thousands) with ages between 15 and 44. IRS Income growth is computed using the adjust gross income from the IRS data. The bottom row indicates which regressions have county effects. All regressions include controls for age and demographic in levels and differences as described in section 1.c. HP is house price growth from Zillow. Demographics are from the Census and the American Community Survey (ACS). Female college is the fraction of women in the county who have at least a bachelor's degree. Old-Households is the instrumental variable that measures fraction of homeowners who are older than 65 years-old and live alone as described in section 1.c. Standard errors are robust and clustered at the state level. ***,**,* coefficient estimate statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; standard errors in parentheses. This table continues in the next page.

	Fertility Change from 2000 to 2006						
	OLS	OLS	OLS	OLS	OLS	IV	1st-stage
Per Cap Origination$_{\{00\to06\}}$	0.29***				0.20***	0.31**	
	(0.04)				(0.04)	(0.13)	
Old-Homeowners$_{\{2000\}}$							22.24***
							(4.09)
HP Growth × Homeownership			15.56***	8.97***	8.90***	8.86***	0.35
			(3.41)	(2.01)	(1.66)	(1.36)	(3.11)
HP Growth$_{\{00\to06\}}$			-4.89***	-1.78	-2.34**	-2.66**	2.78***
			(1.45)	(1.00)	(0.95)	(1.15)	(0.53)
Homeownership$_{\{2000\}}$			-24.94***	-14.22***	-13.37**	-12.88***	-3.45
			(3.52)	(2.84)	(4.22)	(4.38)	(7.93)
House Units Growth$_{\{00\to06\}}$		-6.22**		-3.94*	-4.00*	-4.04***	1.48
		(2.65)		(2.06)	(1.74)	(1.41)	(1.79)
Employment Change$_{\{00\to06\}}$				3.74***	3.87***	3.94***	-0.52
				(0.77)	(0.78)	(0.73)	(0.49)
Female Unemployment$_{\{00\to06\}}$				9.38	9.63	9.77	-2.31
				(11.17)	(9.71)	(8.12)	(7.48)
Male Unemployment$_{\{00\to06\}}$				-23.07**	-23.06**	-23.05***	0.37
				(9.17)	(8.92)	(7.99)	(2.59)
Per Capita Inc. Growth$_{\{00\to06\}}$				0.27	0.11	0.01	0.75
				(2.74)	(2.46)	(2.08)	(1.84)
Log Per Capita Income$_{\{2000\}}$				-6.20**	-5.55**	-5.18***	-2.12
				(1.98)	(1.82)	(1.65)	(1.90)
Log Capital Gains$_{\{2006\}}$				0.30	0.19	0.12	0.20
				(0.66)	(0.67)	(0.62)	(0.25)
Female College (25-44)$_{\{00\to06\}}$				-63.43***	-65.59***	-66.83***	3.90
				(15.21)	(15.29)	(14.00)	(3.56)
Fertility$_{\{2000\}}$				-0.21***	-0.20***	-0.20***	-0.01
				(0.02)	(0.02)	(0.02)	(0.01)
Log Population$_{\{2000\}}$				-2.12***	-2.10***	-2.09***	-0.01
				(0.49)	(0.46)	(0.40)	(0.25)
County Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age-Demo Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Zip Codes	2753	2753	2753	2753	2753	2753	2753
R-squared	0.194	0.181	0.198	0.302	0.309	0.307	0.449

Table 7: Mortgage Origination and Fertility During Boom: Zip Code Level - Old-Homeowners *(Continuation)*

This table is the continuation from the table in the previous page. Fem(a,b) is the number of females (in thousands) in the zip code with ages between a and b. Demographics are from the Census and the American Community Survey (ACS). Fertility rate is defined as the number of births divided by the number of women with ages between 15 and 44. Standard errors are robust and clustered at the state level. ***,**,* coefficient estimate statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; standard errors in parentheses.

	Fertility Change from 2000 to 2006						
	OLS	OLS	OLS	OLS	OLS	IV	1st-stage
Fraction of Hisp.$_{\{00\rightarrow06\}}$				21.05**	17.26**	15.10**	17.24***
				(6.63)	(6.84)	(6.64)	(2.97)
Fraction of Black.$_{\{00\rightarrow06\}}$				20.85**	16.77*	14.44*	19.51***
				(8.62)	(8.31)	(8.12)	(5.28)
Fem (15,17)/Fem (15,45)$_{\{00\rightarrow06\}}$				-48.88**	-53.50**	-56.14***	25.41***
				(20.62)	(20.43)	(19.29)	(7.47)
Fem (18,24)/Fem (15,45)$_{\{00\rightarrow06\}}$				-61.14***	-62.53***	-63.32***	5.39
				(8.15)	(7.96)	(7.36)	(2.91)
Fem (25,34)/Fem (15,45)$_{\{00\rightarrow06\}}$				7.76	5.37	4.01	12.06**
				(8.94)	(8.96)	(8.55)	(3.70)
Fem (15,17)/Fem (15,45)$_{\{2000\}}$				134.33***	115.88***	105.34***	107.78***
				(36.17)	(33.97)	(33.26)	(27.80)
Fem (18,24)/Fem (15,45)$_{\{2000\}}$				16.37*	12.21	9.83	23.45**
				(8.05)	(6.88)	(7.11)	(10.13)
Fem (25,34)/Fem (15,45)$_{\{2000\}}$				58.28***	51.10***	47.00***	40.51*
				(14.28)	(13.10)	(14.05)	(19.72)
County Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age-Demo Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Zip Codes	2753	2753	2753	2753	2753	2753	2753
R-squared	0.194	0.181	0.198	0.302	0.309	0.307	0.449

Table 8: The Net House Wealth Effect: Zip-Code Level

Old-Homeowners is the instrumental variable that measures fraction of homeowners who are older than 65 years-old and live alone as described in section 1.c. Per Cap Origination is defined as the mortgage origination for home purchase divided by the population (in thousands) in the zip code. Fertility rate is defined as the number of births divided by the number of women (in thousands) with ages between 15 and 44. IRS Income growth is computed using the adjust gross income from the IRS data. The bottom row indicates which regressions have county effects. All regressions include controls for age and demographic in levels and differences as described in section 1.c. HP is house price growth from Zillow. Female college is the fraction of women in the county who have at least a bachelor's degree. Standard errors are robust and clustered at the state level. ***,**,* coefficient estimate statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; standard errors in parentheses.

	Fertility Change from 2000 to 2006					
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) IV
Per Cap Origination$_{\{00\to06\}}$	0.29***				0.20***	0.31**
	(0.04)				(0.04)	(0.14)
Old-Homeowners$_{\{2000\}}$						
HP Growth × Homeownership		15.56***				
		(3.41)				
HP Growth$_{\{00\to06\}}$		-4.89***	4.27***	2.70***	2.10**	1.77**
		(1.45)	(1.08)	(0.80)	(0.65)	(0.82)
Homeownership$_{\{2000\}}$		-24.94***		-2.74	-1.97	-1.54
		(3.52)		(4.78)	(5.64)	(5.42)
House Units Growth$_{\{00\to06\}}$				-3.98*	-4.04*	-4.08***
				(2.10)	(1.78)	(1.44)
Employment Change$_{\{00\to06\}}$				3.78***	3.91***	3.98***
				(0.84)	(0.85)	(0.79)
Female Unemployment$_{\{00\to06\}}$				9.40	9.64	9.78
				(10.78)	(9.31)	(7.77)
Male Unemployment$_{\{00\to06\}}$				-20.39*	-20.41*	-20.41**
				(9.53)	(9.23)	(8.25)
Per Capita Inc. Growth$_{\{00\to06\}}$				0.47	0.30	0.21
				(2.54)	(2.26)	(1.92)
Log Per Capita Income$_{\{2000\}}$				-6.77***	-6.11***	-5.75***
				(1.87)	(1.75)	(1.57)
Log Capital Gains$_{\{2006\}}$				0.42	0.30	0.24
				(0.65)	(0.67)	(0.61)
Female College (25-44)$_{\{00\to06\}}$				-63.16***	-65.33***	-66.55***
				(15.39)	(15.47)	(14.21)
Fertility$_{\{2000\}}$				-0.22***	-0.21***	-0.21***
				(0.02)	(0.02)	(0.02)
Log Population$_{\{2000\}}$				-2.17***	-2.15***	-2.14***
				(0.53)	(0.49)	(0.43)
County Effects	Yes	Yes	Yes	Yes	Yes	Yes
Age Controls	Yes	Yes	Yes	Yes	Yes	Yes
#Zip Codes	2753	2753	2753	2753	2753	2753
R-squared	0.194	0.198	0.179	0.298	0.306	0.304

Table 9: Mortgage Origination and Fertility During Boom: Zip Code Level - 75-Homeowners

75-Homeowners is the instrumental variable that measures fraction of homeowners who are older than 75 years-old and live alone as described in section 1.c. Per Cap Origination is defined as the mortgage origination for home purchase divided by the population (in thousands) in the county. Fertility rate is defined as the number of births divided by the number of women (in thousands) with ages between 15 and 44. IRS Income growth is computed using the adjust gross income from the IRS data. The bottom row indicates which regressions have county effects. All regressions include controls for age and demographic in levels and differences as described in section 1.c. HP is house price growth from zillow. Female college is the fraction of women in the county who have at least a bachelor's degree. Standard errors are robust and clustered at the state level. ***,**,* coefficient estimate statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; standard errors in parentheses.

	Fertility Change from 2000 to 2006						
	OLS	OLS	OLS	OLS	OLS	IV	1st-stage
Per Cap Origination$_{\{00\to06\}}$	0.29***				0.20***	0.31**	
	(0.04)				(0.04)	(0.16)	
75-Homeowners$_{\{2000\}}$							23.16***
							(4.15)
HP Growth × Homeownership			15.56***	8.97***	8.90***	8.86***	0.26
			(3.41)	(2.01)	(1.66)	(1.37)	(3.15)
HP Growth$_{\{00\to06\}}$			-4.89***	-1.78	-2.34**	-2.66**	2.87***
			(1.45)	(1.00)	(0.95)	(1.10)	(0.55)
Homeownership$_{\{2000\}}$			-24.94***	-14.22***	-13.37**	-12.87***	-3.79
			(3.52)	(2.84)	(4.22)	(4.73)	(8.18)
House Units Growth$_{\{00\to06\}}$		-6.22**		-3.94*	-4.00*	-4.04***	1.13
		(2.65)		(2.06)	(1.74)	(1.42)	(1.82)
Employment Change$_{\{00\to06\}}$				3.74***	3.87***	3.94***	-0.59
				(0.77)	(0.78)	(0.72)	(0.47)
Female Unemployment$_{\{00\to06\}}$				9.38	9.63	9.77	-1.82
				(11.17)	(9.71)	(8.05)	(7.49)
Male Unemployment$_{\{00\to06\}}$				-23.07**	-23.06**	-23.05***	-0.12
				(9.17)	(8.92)	(7.98)	(2.52)
Per Capita Inc. Growth$_{\{00\to06\}}$				0.27	0.11	0.01	0.66
				(2.74)	(2.46)	(2.09)	(1.82)
Log Per Capita Income$_{\{2000\}}$				-6.20**	-5.55**	-5.18***	-2.45
				(1.98)	(1.82)	(1.68)	(1.83)
Log Capital Gains$_{\{2006\}}$				0.30	0.19	0.12	0.31
				(0.66)	(0.67)	(0.63)	(0.26)
Female College (25-44)$_{\{00\to06\}}$				-63.43***	-65.59***	-66.85***	6.14
				(15.21)	(15.29)	(14.75)	(3.36)
Fertility$_{\{2000\}}$				-0.21***	-0.20***	-0.20***	-0.01
				(0.02)	(0.02)	(0.02)	(0.01)
Log Population$_{\{2000\}}$				-2.12***	-2.10***	-2.09***	-0.06
				(0.49)	(0.46)	(0.39)	(0.26)
County Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age-Demo Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Zip Codes	2753	2753	2753	2753	2753	2753	2753
R-squared	0.194	0.181	0.198	0.302	0.309	0.307	0.446

Table 10: Mortgage Origination and Fertility in Normal Times: Zip-Code Level

Per Cap Origination is defined as the mortgage origination for home purchase divided by the population (in thousands) in the zip code. Fertility rate is defined as the number of births divided by the number of women (in thousands) with ages between 15 and 44. The third column reports the same correlation than column two but in the sub-sample that matches the sample of the first column. The bottom row indicates which regressions have county effects. Standard errors are robust and clustered at the state level. ***,**,* coefficient estimate statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; standard errors in parentheses.

	Fertility Change		
	1995-2000	2000-2006	2000-2006
Per Cap Origination$_{\{95\to00\}}$	-0.07		
	(0.06)		
Per Cap Origination$_{\{00\to06\}}$		0.29***	0.27***
		(0.04)	(0.05)
County Effects	Yes	Yes	Yes
Age-Demo Controls	Yes	Yes	Yes
#Zip Codes	2015	2753	2015
R-squared	0.216	0.194	0.201

Table 11: Change in Homeownership and New Construction: Zip-Code Level

Each column reports the correlation coefficient of the zip code growth in housing units with the change in homeownership from 2000 to 2006 by age group. Construction growth is defined as the growth in number of housing units in the zip code. The bottom row indicates which regressions have county effects. Standard errors are robust and clustered at the state level. ***,**,* coefficient estimate statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; standard errors in parentheses.

	Ownership Change from 2000 to 2006		
	(1)	(2)	(3)
	25< age <44	45< age <59	60< age <74
House Units Growth$_{\{00\to06\}}$	-0.01	-0.08***	0.10***
	(-0.43)	(-3.59)	(4.80)
County Effects	Yes	Yes	Yes
#Zip Codes	2753	2753	2753
R-squared	0.173	0.126	0.166

Table 12: Fertility and Change in Homeownership by Income Groups: Zip-Code Level

Each column reports the correlation coefficient of the zip code fertility change between 2000 and 2006 with the change in homeownership from 2000 to 2006 by income group. Income is measured as the household income from the Census and American Community Survey. Fertility rate is defined as the number of births divided by the number of women (in thousands) with ages between 15 and 44. The bottom row indicates which regressions have county effects. Standard errors are robust and clustered at the state level. ***,**,* coefficient estimate statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; standard errors in parentheses.

	Ownership Change from 2000 to 2006				
	(1)	(2)	(3)	(4)	(5)
	Inc <25	25< Inc <50	50< Inc <100	100< Inc <150	Inc >150
$\text{Fertility}_{\{00\to06\}}$	-0.02	0.06***	0.15***	-0.01	-0.13***
	(-1.15)	(3.20)	(8.30)	(-0.63)	(-7.18)
County Effects	Yes	Yes	Yes	Yes	Yes
#Zip Codes	2753	2753	2753	2753	2753
R-squared	0.353	0.237	0.315	0.224	0.320

Table 13: Gains in Homeownership by Income Groups and Old-Homeowners: Zip-Code Level

Each column reports the correlation coefficient of the zip code fraction of old-homeowners in 2000 with the change in homeownership from 2000 to 2006 by income group. Income is measured as the household income from the Census and American Community Survey. Old-Homeowners is the instrumental variable that measures fraction of homeowners who are older than 65 years-old and live alone as described in section 1.c. The bottom row indicates which regressions have county effects. Standard errors are robust and clustered at the state level. ***,**,* coefficient estimate statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; standard errors in parentheses.

	Ownership Change from 2000 to 2006				
	(1)	(2)	(3)	(4)	(5)
	Inc <25	25< Inc <50	50< Inc <100	100< Inc <150	Inc >150
Old Homeowners	-0.15***	0.15***	0.30***	-0.12***	-0.21***
	(-8.09)	(7.82)	(16.65)	(-5.78)	(-11.06)
County Effects	Yes	Yes	Yes	Yes	Yes
#Zip Codes	2753	2753	2753	2753	2753
R-squared	0.368	0.251	0.365	0.233	0.338

Table 14: Women in the Labor Force: Family Level

Mortgage and baby (1) is equal to one if a woman in a household who had a baby during the housing boom and got a mortgage loan, and is equal to zero for a woman in a household who had a baby during the housing boom but was always a renter. *Mortgage and baby (2)* is equal to one if a woman in a household who had a baby during the housing boom and got a mortgage loan, and is equal to zero for a woman in any household. PUMA are geographical divisions from the census that have on average 100,000 people. Year built is the year when the structure where the household lives was built. ***,**,* coefficient estimate statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; t-statistics in parentheses.

	In Labor Force	Employed
	(1)	(2)
	2007≤ t ≤2011	2007≤ t ≤2011
Mortgage and Baby (1)	0.140***	
	(13.97)	
Mortgage and baby (2)		-0.013***
		(-7.39)
Husband Age	-0.000	0.010***
	(-0.35)	(50.94)
Husband Income	-0.000***	0.000***
	(-21.63)	(11.97)
Husband has College	0.004*	0.006***
	(1.95)	(23.28)
Wife Age	0.002**	-0.012***
	(2.34)	(-41.33)
Wife has College	0.134***	-0.029***
	(17.05)	(-12.94)
Year Built	0.005***	-0.001***
	(3.70)	(-6.74)
Number Bedrooms	0.011**	0.008***
	(2.31)	(22.69)
PUMA Effects	Yes	Yes
Race Effects	Yes	Yes
#Households	21472	1580951
R-squared	0.126	0.283

Figure 7: House Prices and House Transactions

The top panel of this figure reports the Federal Housing Finance Agency US House Price Index from 1990 to 2013. The points plotted are the December indexes. The bottom panel depicts the annual volume of homes sales, the green line reports the annual number of house transactions for newly constructed homes and the orange line reports the annual number of house transactions for existing homes. Both levels are presented in millions. The data is from the National Association of Realtors website. The vertical dashed lines highlight the end of year of 2005 and 2006.

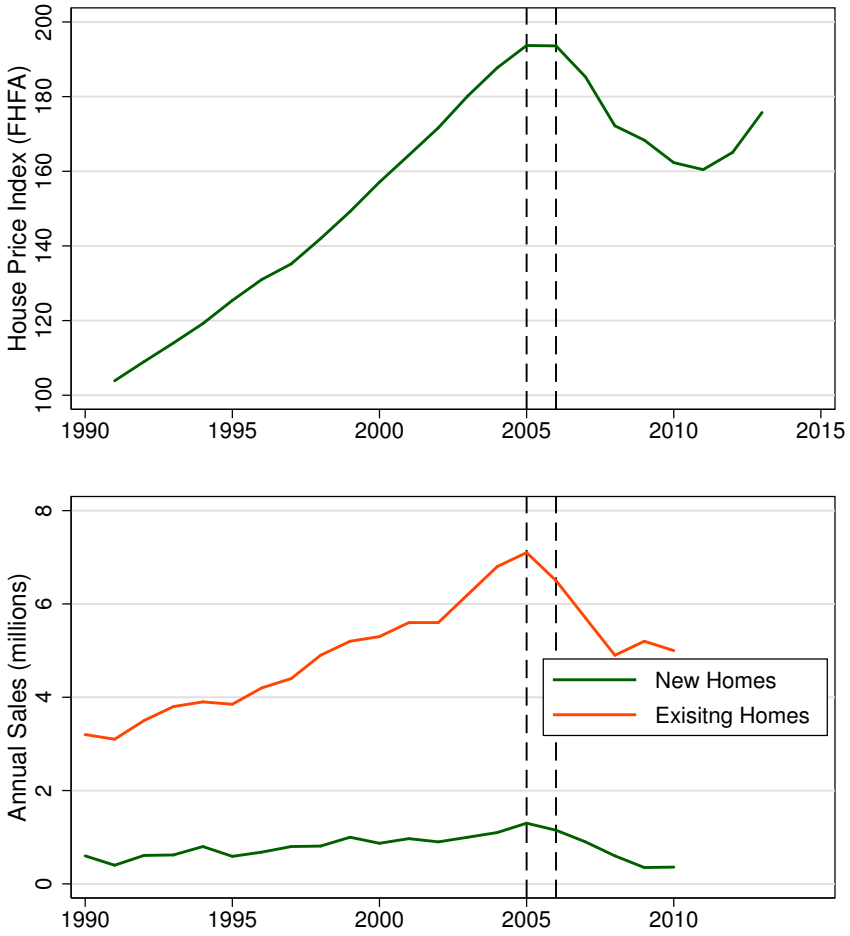

Figure 8: Mortgage Interest Rates from 1999 to 2013

In the chart below, the orange line depicts the 1-year adjustable rate from Fannie Mae and Freddie Mac webpage, while the green line outlines the 30-year fixed rate from the same source. The vertical dashed lines highlight the end of year of 2004 and 2006.

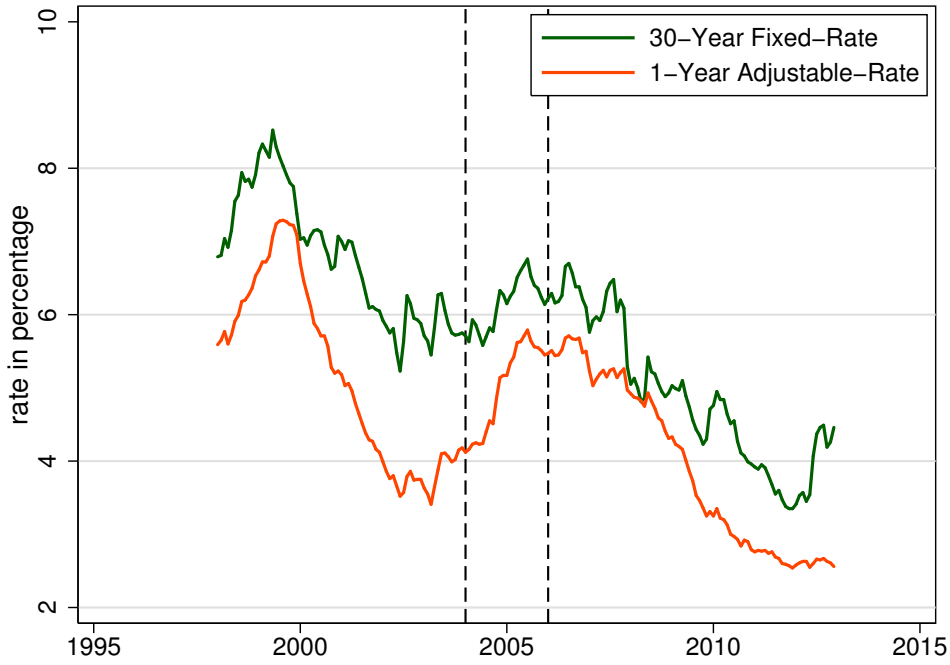

Figure 9: Mortgage Delinquency rates from 1991 to 2013

The chart below plots the delinquency rate on single-family residential Mortgages for the US, booked in domestic offices of all commercial banks. The vertical dashed lines highlight the end of year of 2005 and 2006. The data is from the Economic Research Section of the Federal Reserve Bank of St. Louis.

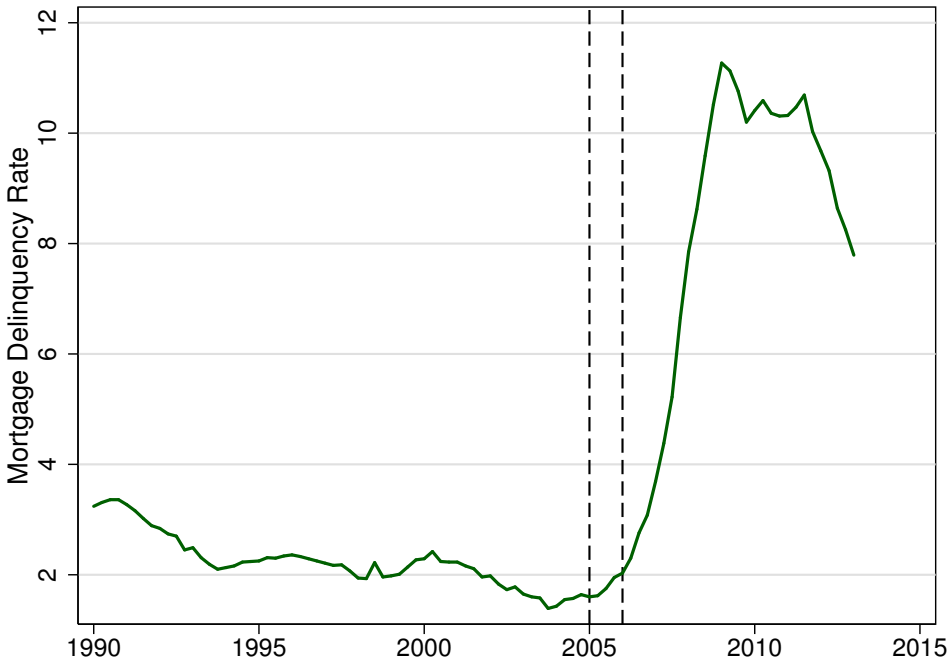

Figure 10: Search Volume for Real Estate Terms

This figure reports the time series of Google search volume for the following search terms:

$$\{real\ estate,\ mls,\ remax,\ century\ 21, realty,\ sothebys,\ era,\ coldwell,\ prudential,\ realtor,\ zillow,\ redfin\}$$

The top left panel outlines the data for California, the top right for Texas, the bottom left for Florida, and the bottom right for Massachusetts. The index plotted is the search volume index (SV) directly downloaded from Google Trends. SVI is explained in detail in section 2.b. For a given geography, it approximately denotes the percentage of total search volume on a set of given search terms. The vertical dashed lines highlight the end of year of 2005 and 2006.

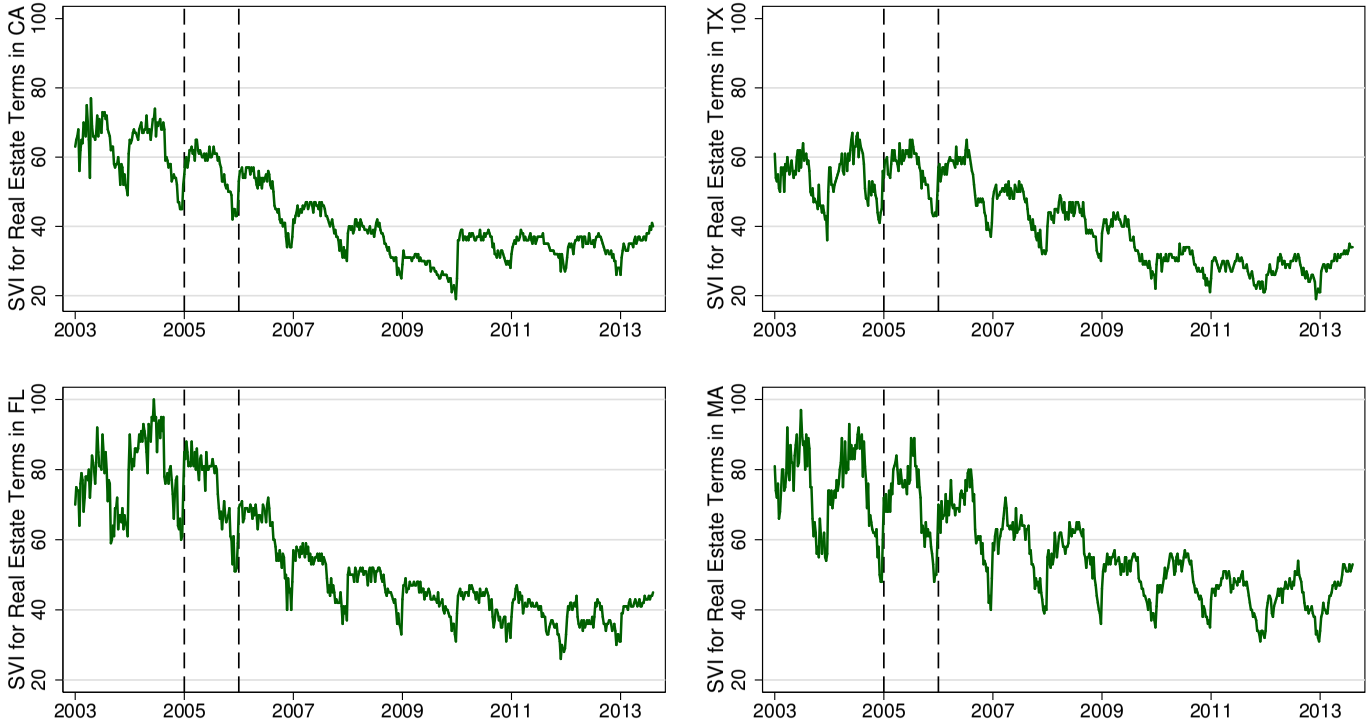

Figure 11: Search Volume for Real Estate Terms versus Housing Outcomes

The figure below shows that changes in SVI are correlated with house price growth and growth in number of mortgage applications. The SVI is computed for the following search terms:

$$\{real\ estate,\ mls,\ remax,\ century\ 21, realty,\ sothebys,\ era,\ coldwell,\ prudential,\ realtor,\ zillow,\ redfin\}$$

SVI is explained in detail in section 2.b. For a given geography, it approximately denotes the percentage of total search volume on a set of given search terms. HP Growth is defined as the growth in the house price index from the Federal Housing Finance Agency. The number of mortgage application is computed from the HMDA dataset by aggregating the applications for home purchase at the county level.

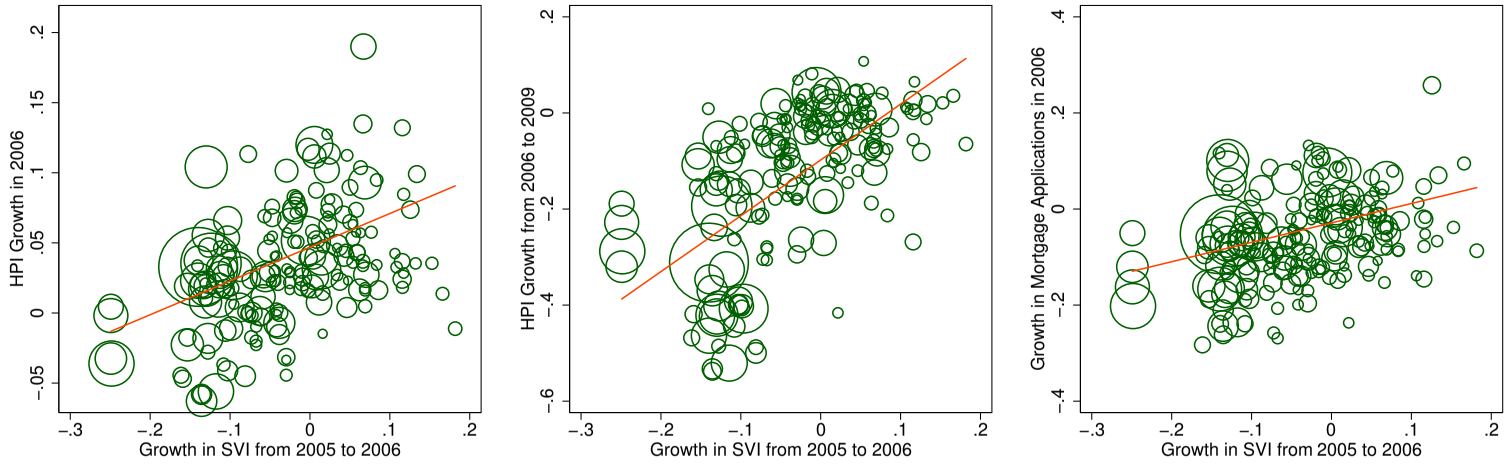

Figure 12: Search Volume for Real Estate Terms versus Instrument

This figure presents a scatter of the instrument described in section 2.c against Google's SVI for the following search terms:

$\{real\ estate,\ mls,\ remax,\ century\ 21, realty,\ sothebys,\ era,\ coldwell,\ prudential,\ realtor,\ zillow,\ redfin\}$

The index plotted is directly downloaded from Google Trends. SVI is explained in detail in section 2.b. For a given geography, it approximately denotes the percentage of total search volume on a set of given search terms.

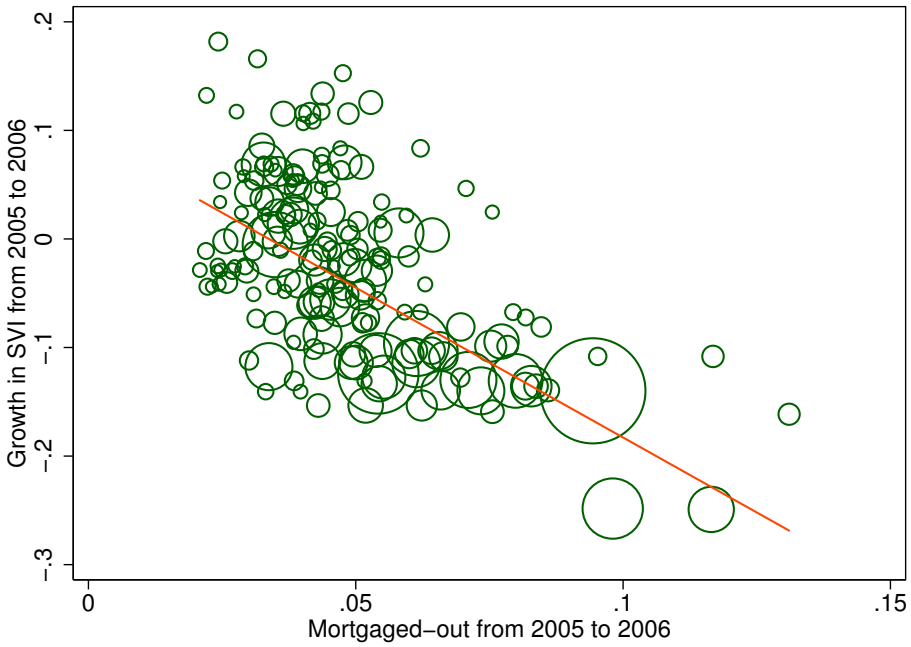

Figure 13: Households Mortgaged-out

The figure below sketches the empirical design proposed in section 2.c. Consider that P is the probability distribution of income, r_{2005} is the interest rate level in 2005, $\Delta r > 0$ is the interest rate change from 2005 to 2006, and $\overline{Loan}$ is the average loan in 2005. The orange (lighter) area, $P(Income < 2 \times r_{2005} \times \overline{Loan}_{2005})$, represents the fraction of the county households who, at the 2005 level of interest rates, are "excluded" from the real estate market because they cannot afford to pay an interest-only mortgage. I posit that a household cannot afford an interest-rate-only mortgage when the mortgage payments are higher than 50% of the household income. The red (darker) area, $P(2 \times r_{2005} \times \overline{Loan}_{2005} < Income < 2 \times (r_{2005} + \Delta r) \times \overline{Loan}_{2005})$, represents the new fraction of households who cannot afford to pay an interest-rate-only mortgage after the increase in interest rates from 2005 to 2006. I define that an household cannot afford an interest-rate-only mortgage, when the mortgage payments are higher than 50% of the household's income.

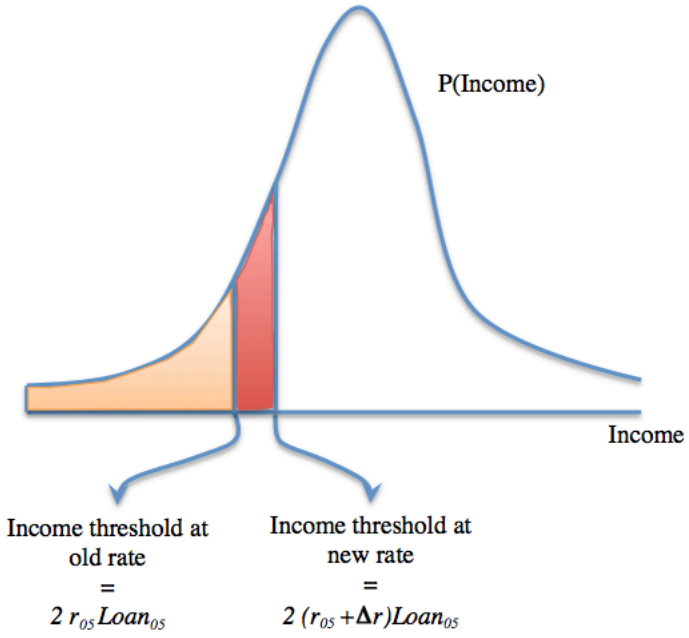

Table 15: Summary of Statistics

$\text{SVI}_{i,\{05\to06\}}$ is defined as the ratio of Google Search Volume Index (SVI) in 2006 over the SVI in 2005. $\text{SVI1}_{\{05\to06\}}$ is the search volume index for the following search terms:

{*real estate, mls, remax, century 21, realty, sothebys, era, coldwell, prudential, realtor, zillow, redfin*}

$\text{SVI2}_{\{05\to06\}}$ is the search volume index for the following search terms:

{*real estate, realtor, realtors, mls, realty*}

$\text{SVI3}_{\{05\to06\}}$ is the search volume index for the following search terms:

{*houses, real estate, construction*}

For a given geography, it approximately denotes the percentage of total search volume on a set of given search terms. *Mortgage Denial Rate*$_{\{05\to06\}}$ is defined as the change in denial rate for all mortgage loans for home purchase. *Mortgage Delinquency*$_{\{05\to06\}}$ is defined as the change in in the percentage of mortgage debt balance that are 90+ days delinquent. *Unemployment Rate*$_{\{05\to06\}}$ is defined as the change in the Bureau of Labor Statistics official unemployment rate. *Income per Capita*$_{\{05\to06\}}$ is defined as the growth in adjusted gross income per capita from the IRS, *FHFA HPI Growth*$_{\{00\to05\}}$ is defined as the growth in the house price index from the Federal Housing Finance Agency. *Mortgaged-out*$_{\{2005\}}$ is defined in detail in section 2.c, and approximately defined as the number of households who cannot afford an interest-only mortgage given the price level in 2005. All statistics are weighted by the population size of the county.

	N	Mean	Std	10th	50th	90th
Mortgaged-out$_{\{05\to06\}}$	670	0.054	0.023	0.034	0.050	0.083
Mortgaged-out$_{\{2005\}}$	670	0.17	0.050	0.12	0.16	0.24
SVI1$_{\{05\to06\}}$	670	0.0019	0.082	-0.10	-0.0052	0.11
SVI2$_{\{05\to06\}}$	670	0.043	0.091	-0.067	0.030	0.17
SVI3$_{\{05\to06\}}$	670	-0.061	0.081	-0.14	-0.073	0.050
Unemployment Rate$_{\{05\to06\}}$	670	-0.49	0.39	-0.90	-0.40	-0.10
Income Per Capita$_{\{05\to06\}}$	670	0.040	0.021	0.016	0.039	0.067
Mortgage Denial Rate$_{\{05\to06\}}$	670	0.017	0.021	-0.0071	0.016	0.047
Mortgage Delinquency$_{\{05\to06\}}$	670	0.31	0.47	-0.14	0.27	0.84
FHFA HPI Growth$_{\{00\to05\}}$	670	0.088	0.048	0.037	0.083	0.17
FHFA HPI Growth$_{\{05\to06\}}$	670	0.034	0.039	-0.0071	0.031	0.090

Table 16: House Price Growth in 2006 and SVI1 for Real Estate Terms: OLS

This table reports the estimates of the following regression model:

$$\text{House Price Growth}_{\ 05\to06,i} = \beta_0 + \beta_1 \times \text{SVI1}_{\ 05\to06,i} + \alpha \times X_i + \varepsilon_i,$$

where i represents a county, $\text{SVI1}_{\{05\to06\}}$ is defined as the ratio of Google Search Volume Index (SVI) in 2006 over the SVI in 2005 for the following search terms:

{*real estate, mls, remax, century 21, realty, sothebys, era, coldwell, prudential, realtor, zillow, redfin*}

SVI is defined in detail in section 2.b. For a given geography, it approximately denotes the percentage of total search volume on a set of given search terms. X is a vector of controls for county economic changes. It controls for the following variables: *Mortgage Denial Rate*$_{\{05\to06\}}$, defined as the change in denial rate for all mortgage loans for home purchase; *Mortgage Delinquency*$_{\{05\to06\}}$, defined as the change in the percentage of mortgage debt balance that are 90+ days delinquent; *Unemployment Rate*$_{\{05\to06\}}$, defined as the change in the Bureau of Labor Statistics official unemployment rate; *Income per Capita*$_{\{05\to06\}}$, defined as the growth in adjusted gross income per capita from the Internal Revenue Service; *FHFA HPI Growth*$_{\{00\to05\}}$, defined as the growth in the house price index from the Federal Housing Finance Agency; *Mortgaged-out*$_{\{2005\}}$, defined in detail in section 2.c, and approximately defined as the fraction of households that cannot afford an interest-only mortgage on the county's average loan in 2005. All variables are standardized. ***,**,* coefficient estimates are statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; t-statistics are in parentheses. Standard errors are heteroskedastically robust, and are clustered at the state level.

	House Price Growth during 2006						
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS
$\text{SVI1}_{\{05\to06\}}$	0.352***		0.358***	0.377***	0.303**	0.336***	0.347***
	(2.89)		(3.94)	(4.45)	(2.59)	(3.20)	(4.30)
$\text{Mortgage Denial Rate}_{\{05\to06\}}$		-0.043	0.001		0.094		0.014
		(-0.45)	(0.01)		(0.80)		(0.12)
$\text{Mortgage Delinquency}_{\{05\to06\}}$		-0.261***	-0.247***		-0.315***		-0.247***
		(-6.07)	(-4.98)		(-4.38)		(-5.01)
$\text{Unemployment Rate}_{\{05\to06\}}$		-0.137**	-0.140**	-0.135*			-0.148**
		(-2.46)	(-2.01)	(-1.69)			(-2.03)
$\text{Income Per Capita}_{\{05\to06\}}$		0.351***	0.342***	0.392***			0.345***
		(5.14)	(4.87)	(5.49)			(4.94)
$\text{FHFA HPI Growth}_{\{00\to05\}}$		0.013	0.184*	0.104			0.253**
		(0.16)	(1.77)	(0.89)			(2.52)
$\text{Mortgaged-out}_{\{2005\}}$						-0.037	-0.132**
						(-0.61)	(-2.55)
#Counties	676	676	676	676	676	676	676
R-squared	0.124	0.286	0.372	0.323	0.205	0.125	0.382

Table 17: House Price Growth in 2006 and SVI1 for Real Estate Terms: IV

This table resorts the estimates of the IV model. $\text{SVI1}_{\{05\to06\}}$ is defined as the ratio of Google Search Volume Index (SVI) in 2006 over the SVI in 2005 for the following search terms:

{*real estate, mls, remax, century 21, realty, sothebys, era, coldwell, prudential, realtor, zillow, redfin*}

SVI is defined in detail in section 2.b. For a given geography, it approximately denotes the percentage of total search volume on a set of given search terms. $\textit{Mortgaged-out}_{\{05\to06\}}$, defined in detail in section 2.c, and approximately defined as the fraction of households that cannot afford an interest-only mortgage on the county's average loan in 2005, after the increase in interest rates from 2005 to 2006. X is a vector of controls for county economic changes. It controls for the following variables: $\textit{Mortgage Denial Rate}_{\{05\to06\}}$, defined as the change in denial rate for all mortgage loans for home purchase; $\textit{Mortgage Delinquency}_{\{05\to06\}}$, defined as the change in the percentage of mortgage debt balance that are 90+ days delinquent; $\textit{Unemployment Rate}_{\{05\to06\}}$, defined as the change in the Bureau of Labor Statistics official unemployment rate; $\textit{Income per Capita}_{\{05\to06\}}$, defined as the growth in adjusted gross income per capita from the Internal Revenue Service; $\textit{FHFA HPI Growth}_{\{00\to05\}}$, defined as the growth in the house price index from the Federal Housing Finance Agency; $\textit{Mortgaged-out}_{\{2005\}}$, defined in detail in section 2.c, and approximately defined as the fraction of households that cannot afford an interest-only mortgage on the county's average loan in 2005. All variables are standardized. ***,**,* coefficient estimates are statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; t-statistics are in parentheses. Standard errors are heteroskedastically robust, and are clustered at the state level.

	$\text{SVI1}_{\{05\to06\}}$		House Price Growth during 2006				
	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV	(7) IV
$\text{SVI1}_{\{05\to06\}}$			0.347***	0.394**	0.447**	0.384**	0.417***
			(4.30)	(2.25)	(2.22)	(2.03)	(2.66)
$\text{Mortgaged-out}_{\{05\to06\}}$	-0.586***	-0.707***					
	(-8.68)	(-8.54)					
$\text{Mortgaged-out}_{\{2005\}}$		0.414***	-0.132**		-0.092	-0.136**	-0.126**
		(3.80)	(-2.55)		(-1.24)	(-2.13)	(-2.18)
$\text{Unemployment Rate}_{\{05\to06\}}$		-0.033	-0.148**			-0.147*	-0.148**
		(-0.72)	(-2.03)			(-1.80)	(-2.03)
$\text{Income Per Capita}_{\{05\to06\}}$		-0.013	0.345***			0.393***	0.343***
		(-0.18)	(4.94)			(5.30)	(4.95)
$\text{FHFA HPI Growth}_{\{00\to05\}}$		-0.262***	0.253**		0.163	0.193	0.283**
		(-3.13)	(2.52)		(1.00)	(1.21)	(2.35)
$\text{Mortgage Denial Rate}_{\{05\to06\}}$		-0.108*	0.014				0.021
		(-1.69)	(0.12)				(0.20)
$\text{Mortgage Delinquency}_{\{05\to06\}}$		-0.057	-0.247***				-0.244***
		(-1.47)	(-5.01)				(-4.59)
#Counties	676	676	676	676	676	676	676
R-squared	0.344	0.447	0.382	0.122	0.131	0.334	0.379

Table 18: House Price Growth in 2006 and SVI2 for Real Estate Terms: OLS

This table reports the estimates of the following regression model:

$$\text{House Price Growth}_{\ 05\to06,i} = \beta_0 + \beta_1 \times \text{SVI2}_{\ 05\to06,i} + \alpha \times X_i + \varepsilon_i,$$

where i represents a county, $\text{SVI2}_{\{05\to06\}}$ is defined as the ratio of Google Search Volume Index (SVI) in 2006 over the SVI in 2005 for the following search terms:

$$\{\textit{real estate, realtor, realtors, mls, realty}\}$$

SVI is defined in detail in section 2.b. For a given geography, it approximately denotes the percentage of total search volume on a set of given search terms. X is a vector of controls for county economic changes. It controls for the following variables: *Mortgage Denial Rate*$_{\{05\to06\}}$, defined as the change in denial rate for all mortgage loans for home purchase; *Mortgage Delinquency*$_{\{05\to06\}}$, defined as the change in the percentage of mortgage debt balance that are 90+ days delinquent; *Unemployment Rate*$_{\{05\to06\}}$, defined as the change in the Bureau of Labor Statistics official unemployment rate; *Income per Capita*$_{\{05\to06\}}$, defined as the growth in adjusted gross income per capita from the Internal Revenue Service; *FHFA HPI Growth*$_{\{00\to05\}}$, defined as the growth in the house price index from the Federal Housing Finance Agency; *Mortgaged-out*$_{\{2005\}}$, defined in detail in section 2.c, and approximately defined as the fraction of households that cannot afford an interest-only mortgage on the county's average loan in 2005. All variables are standardized. ***,**,* coefficient estimates are statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; t-statistics are in parentheses. Standard errors are heteroskedastically robust, and are clustered at the state level.

	House Price Growth during 2006						
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS
$\text{SVI2}_{\{05\to06\}}$	0.384***		0.415***	0.438***	0.355***	0.381***	0.405***
	(3.19)		(5.38)	(5.43)	(3.32)	(3.50)	(5.57)
$\text{Mortgage Denial Rate}_{\{05\to06\}}$		-0.047	0.023		0.146		0.036
		(-0.50)	(0.22)		(1.29)		(0.34)
$\text{Mortgage Delinquency}_{\{05\to06\}}$		-0.256***	-0.235***		-0.304***		-0.235***
		(-6.00)	(-4.68)		(-4.10)		(-4.63)
$\text{Unemployment Rate}_{\{05\to06\}}$		-0.144**	-0.189***	-0.193***			-0.195***
		(-2.55)	(-3.03)	(-2.97)			(-3.08)
$\text{Income Per Capita}_{\{05\to06\}}$		0.351***	0.298***	0.344***			0.302***
		(5.09)	(4.06)	(4.62)			(4.11)
$\text{FHFA HPI Growth}_{\{00\to05\}}$		0.013	0.213*	0.153			0.286***
		(0.17)	(1.94)	(1.17)			(2.82)
$\text{Mortgaged-out}_{\{2005\}}$						-0.008	-0.137***
						(-0.13)	(-2.89)
#Counties	670	670	670	670	670	670	670
R-squared	0.148	0.288	0.385	0.343	0.223	0.148	0.396

Table 19: House Price Growth in 2006 and SVI2 for Real Estate Terms: IV

This table reports the estimates of the IV model. SVI is defined in detail in section 2.b. For a given geography, it approximately denotes the percentage of total search volume on a set of given search terms. *Mortgaged-out*$_{\{05\to06\}}$, defined in detail in section 2.c, and approximately defined as the fraction of households that cannot afford an interest-only mortgage on the county's average loan in 2005, after the increase in interest rates from 2005 to 2006. X is a vector of controls for county economic changes. It controls for the following variables: *Mortgage Denial Rate*$_{\{05\to06\}}$, defined as the change in denial rate for all mortgage loans for home purchase; *Mortgage Delinquency*$_{\{05\to06\}}$, defined as the change in the percentage of mortgage debt balance that are 90+ days delinquent; *Unemployment Rate*$_{\{05\to06\}}$, defined as the change in the Bureau of Labor Statistics official unemployment rate; *Income per Capita*$_{\{05\to06\}}$, defined as the growth in adjusted gross income per capita from the Internal Revenue Service; *FHFA HPI Growth*$_{\{00\to05\}}$, defined as the growth in the house price index from the Federal Housing Finance Agency; *Mortgaged-out*$_{\{2005\}}$, defined in detail in section 2.c, and approximately defined as the fraction of households that cannot afford an interest-only mortgage on the county's average loan in 2005. All variables are standardized. ***,**,* coefficient estimates are statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; t-statistics are in parentheses. Standard errors are heteroskedastically robust, and are clustered at the state level.

	SVI2 $_{\{05\to06\}}$		House Price Growth during 2006				
	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV	(7) IV
SVI2$_{\{05\to06\}}$			0.405***	0.358**	0.377**	0.343**	0.373***
			(5.57)	(2.14)	(2.42)	(2.16)	(2.88)
Mortgaged-out$_{\{05\to06\}}$	-0.644***	-0.801***					
	(-6.17)	(-10.63)					
Mortgaged-out$_{\{2005\}}$		0.508***	-0.137***		-0.105*	-0.148***	-0.139***
		(5.96)	(-2.89)		(-1.73)	(-2.66)	(-2.67)
Unemployment Rate$_{\{05\to06\}}$		0.062	-0.195***			-0.191***	-0.192***
		(1.19)	(-3.08)			(-3.10)	(-3.44)
Income Per Capita$_{\{05\to06\}}$		0.084*	0.302***			0.360***	0.306***
		(1.73)	(4.11)			(4.16)	(4.02)
FHFA HPI Growth$_{\{00\to05\}}$		-0.259***	0.286***		0.151	0.189	0.272**
		(-3.74)	(2.82)		(0.89)	(1.13)	(2.17)
Mortgage Denial Rate$_{\{05\to06\}}$		-0.158***	0.036				0.031
		(-2.77)	(0.34)				(0.28)
Mortgage Delinquency$_{\{05\to06\}}$		-0.071**	-0.235***				-0.236***
		(-2.05)	(-4.63)				(-4.52)
#Counties	670	670	670	670	670	670	670
R-squared	0.414	0.576	0.396	0.147	0.162	0.351	0.396

Table 20: House Price Growth in 2006 and SVI3 for Real Estate Terms: OLS

This table reports the estimates of the following regression model:

$$\text{House Price Growth}_{\ 05\to06,i} = \beta_0 + \beta_1 \times \text{SVI3}_{\ 05\to06,i} + \alpha \times X_i + \varepsilon_i,$$

where i represents a county, $\text{SVI3}_{\{05\to06\}}$ is defined as the ratio of Google Search Volume Index (SVI) in 2006 over the SVI in 2005 for the following search terms:

$$\{\textit{houses, real estate, construction}\}$$

SVI is defined in detail in section 2.b. For a given geography, it approximately denotes the percentage of total search volume on a set of given search terms. X is a vector of controls for county economic changes. It controls for the following variables: *Mortgage Denial Rate*$_{\{05\to06\}}$, defined as the change in denial rate for all mortgage loans for home purchase; *Mortgage Delinquency*$_{\{05\to06\}}$, defined as the change in the percentage of mortgage debt balance that are 90+ days delinquent; *Unemployment Rate*$_{\{05\to06\}}$, defined as the change in the Bureau of Labor Statistics official unemployment rate; *Income per Capita*$_{\{05\to06\}}$, defined as the growth in adjusted gross income per capita from the Internal Revenue Service; *FHFA HPI Growth*$_{\{00\to05\}}$, defined as the growth in the house price index from the Federal Housing Finance Agency; *Mortgaged-out*$_{\{2005\}}$, defined in detail in section 2.c, and approximately defined as the fraction of households that cannot afford an interest-only mortgage on the county's average loan in 2005. All variables are standardized. ***,**,* coefficient estimates are statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; t-statistics are in parentheses. Standard errors are heteroskedastically robust, and are clustered at the state level.

	House Price Growth during 2006						
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS
$\text{SVI3}_{\{05\to06\}}$	0.409***		0.446***	0.475***	0.381***	0.439***	0.429***
	(4.34)		(5.91)	(6.27)	(3.80)	(4.31)	(5.67)
$\text{Mortgage Denial Rate}_{\{05\to06\}}$		-0.054	0.015		0.144		0.021
		(-0.58)	(0.14)		(1.24)		(0.20)
$\text{Mortgage Delinquency}_{\{05\to06\}}$		-0.260***	-0.236***		-0.287***		-0.236***
		(-6.19)	(-4.67)		(-4.31)		(-4.63)
$\text{Unemployment Rate}_{\{05\to06\}}$		-0.139**	-0.160***	-0.160**			-0.165***
		(-2.57)	(-2.69)	(-2.64)			(-2.79)
$\text{Income Per Capita}_{\{05\to06\}}$		0.348***	0.285***	0.328***			0.290***
		(5.18)	(4.41)	(5.25)			(4.45)
$\text{FHFA HPI Growth}_{\{00\to05\}}$		0.021	0.256**	0.197			0.295**
		(0.28)	(2.28)	(1.53)			(2.67)
$\text{Mortgaged-out}_{\{2005\}}$						0.056	-0.086
						(0.81)	(-1.63)
#Counties	692	692	692	692	692	692	692
R-squared	0.167	0.286	0.391	0.347	0.235	0.170	0.395

Table 21: House Price Growth in 2006 and SVI3 for Real Estate Terms: IV

This table reports the estimates of the IV model. SVI is defined in detail in section 2.b. For a given geography, it approximately denotes the percentage of total search volume on a set of given search terms. *Mortgaged-out*$_{\{05\to06\}}$, defined in detail in section 2.c, and approximately defined as the fraction of households that cannot afford an interest-only mortgage on the county's average loan in 2005, after the increase in interest rates from 2005 to 2006. X is a vector of controls for county economic changes. It controls for the following variables: *Mortgage Denial Rate*$_{\{05\to06\}}$, defined as the change in denial rate for all mortgage loans for home purchase; *Mortgage Delinquency*$_{\{05\to06\}}$, defined as the change in the percentage of mortgage debt balance that are 90+ days delinquent; *Unemployment Rate*$_{\{05\to06\}}$, defined as the change in the Bureau of Labor Statistics official unemployment rate; *Income per Capita*$_{\{05\to06\}}$, defined as the growth in adjusted gross income per capita from the Internal Revenue Service; *FHFA HPI Growth*$_{\{00\to05\}}$, defined as the growth in the house price index from the Federal Housing Finance Agency; *Mortgaged-out*$_{\{2005\}}$, defined in detail in section 2.c, and approximately defined as the fraction of households that cannot afford an interest-only mortgage on the county's average loan in 2005. All variables are standardized. ***,**,* coefficient estimates are statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; t-statistics are in parentheses. Standard errors are heteroskedastically robust, and are clustered at the state level.

	SVI3 $_{\{05\to06\}}$		House Price Growth during 2006				
	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV	(7) IV
SVI3$_{\{05\to06\}}$			0.429***	0.334**	0.467**	0.419*	0.459**
			(5.67)	(2.30)	(2.30)	(1.91)	(2.51)
Mortgaged-out$_{\{05\to06\}}$	-0.672***	-0.623***					
	(-8.85)	(-7.61)					
Mortgaged-out$_{\{2005\}}$		0.267***	-0.086		-0.049	-0.095	-0.081
		(3.49)	(-1.63)		(-0.59)	(-1.12)	(-1.10)
Unemployment Rate$_{\{05\to06\}}$		0.008	-0.165***			-0.164***	-0.166***
		(0.09)	(-2.79)			(-3.01)	(-2.96)
Income Per Capita$_{\{05\to06\}}$		0.110**	0.290***			0.338***	0.285***
		(2.23)	(4.45)			(3.89)	(3.88)
FHFA HPI Growth$_{\{00\to05\}}$		-0.277***	0.295**		0.192	0.222	0.308**
		(-3.69)	(2.67)		(1.12)	(1.26)	(2.20)
Mortgage Denial Rate$_{\{05\to06\}}$		-0.130*	0.021				0.025
		(-1.76)	(0.20)				(0.25)
Mortgage Delinquency$_{\{05\to06\}}$		-0.073*	-0.236***				-0.234***
		(-1.99)	(-4.63)				(-4.14)
#Counties	692	692	692	692	692	692	692
R-squared	0.451	0.574	0.395	0.162	0.191	0.351	0.395

Table 22: House Price Growth in the Crisis and SVI for Real Estate Terms

This table reports the estimates for the OLS and IV regression model for the crisis. SVI is defined in detail in section 2.b. For a given geography, it approximately denotes the percentage of total search volume on a set of given search terms. *Mortgaged-out*$_{\{05\to06\}}$, defined in detail in section 2.c, and approximately defined as the fraction of households that cannot afford an interest-only mortgage on the county's average loan in 2005, after the increase in interest rates from 2005 to 2006. X is a vector of controls for county economic changes. It controls for the following variables: *Mortgage Denial Rate*$_{\{05\to06\}}$, defined as the change in denial rate for all mortgage loans for home purchase; *Mortgage Delinquency*$_{\{05\to06\}}$, defined as the change in the percentage of mortgage debt balance that are 90+ days delinquent; *Unemployment Rate*$_{\{05\to06\}}$, defined as the change in the Bureau of Labor Statistics official unemployment rate; *Income per Capita*$_{\{05\to06\}}$, defined as the growth in adjusted gross income per capita from the Internal Revenue Service; *Mortgaged-out*$_{\{2005\}}$, defined in detail in section 2.c, and approximately defined as the fraction of households that cannot afford an interest-only mortgage on the county's average loan in 2005. All variables are standardized. ***,**,* coefficient estimates are statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; t-statistics are in parentheses. Standard errors are heteroskedastically robust, and are clustered at the state level.

	House Price Growth form 2006 to 2009					
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) IV	(6) IV
SVI1$_{\{05\to06\}}$	0.563***		0.309***	0.273***	0.662***	0.603***
	(5.93)		(6.40)	(5.54)	(3.18)	(2.80)
Mortgage Denial Rate$_{\{05\to06\}}$		-0.534***	-0.418***	-0.379***	-0.284***	-0.292***
		(-11.82)	(-6.33)	(-5.29)	(-2.93)	(-3.38)
Mortgage Delinquency$_{\{05\to06\}}$		-0.262***	-0.224***	-0.216***	-0.182**	-0.185**
		(-4.95)	(-3.53)	(-3.31)	(-2.26)	(-2.32)
Unemployment Rate$_{\{05\to06\}}$		-0.068	-0.055	-0.056*	-0.040	-0.042
		(-1.42)	(-1.54)	(-1.69)	(-1.08)	(-1.25)
Income Per Capita$_{\{05\to06\}}$		0.184***	0.182***	0.188***	0.180***	0.182***
		(4.35)	(4.37)	(4.65)	(3.46)	(3.60)
Mortgaged-out$_{\{2005\}}$				-0.128**		-0.034
				(-2.15)		(-0.44)
#Counties	676	676	676	676	676	676
R-squared	0.317	0.552	0.629	0.641	0.528	0.560

Table 23: House Price Growth in 2006 and SVI1 for Real Estate Terms: Robustness

This table reports the estimates for the OLS and IV model when house price level in 2005 is used as control, instead of house price growth from 2000 to 2005. SVI is defined in detail in section 2.b. For a given geography, it approximately denotes the percentage of total search volume on a set of given search terms. *Mortgaged-out*$_{\{05\to06\}}$, defined in detail in section 2.c, and approximately defined as the fraction of households that cannot afford an interest-only mortgage on the county's average loan in 2005, after the increase in interest rates from 2005 to 2006. X is a vector of controls for county economic changes. It controls for the following variables: *Mortgage Denial Rate*$_{\{05\to06\}}$, defined as the change in denial rate for all mortgage loans for home purchase; *Mortgage Delinquency*$_{\{05\to06\}}$, defined as the change in the percentage of mortgage debt balance that are 90+ days delinquent; *Unemployment Rate*$_{\{05\to06\}}$, defined as the change in the Bureau of Labor Statistics official unemployment rate; *Income per Capita*$_{\{05\to06\}}$, defined as the growth in adjusted gross income per capita from the Internal Revenue Service; *FHFA HPI*$_{\{2005\}}$, defined as the growth in the house price index from the Federal Housing Finance Agency; *Mortgaged-out*$_{\{2005\}}$, defined in detail in section 2.c, and approximately defined as the fraction of households that cannot afford an interest-only mortgage on the county's average loan in 2005. All variables are standardized. ***,**,* coefficient estimates are statistically distinct from 0 at the 1%, 5% and 10% levels, respectively; t-statistics are in parentheses. Standard errors are heteroskedastically robust, and are clustered at the state level.

	SVI1 $_{\{05\to06\}}$		House Price Growth during 2006			
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) IV	(6) IV
SVI1$_{\{05\to06\}}$			0.352***	0.332***	0.394**	0.454**
			(2.89)	(3.67)	(2.25)	(2.42)
Mortgaged-out$_{\{05\to06\}}$	-0.586***	-0.604***				
	(-8.68)	(-7.52)				
Mortgaged-out$_{\{2005\}}$		0.395***		-0.116		-0.120
		(3.54)		(-1.47)		(-1.61)
Unemployment Rate$_{\{05\to06\}}$		-0.026		-0.138*		-0.140*
		(-0.65)		(-1.86)		(-1.90)
Income Per Capita$_{\{05\to06\}}$		-0.011		0.349***		0.347***
		(-0.15)		(4.91)		(4.92)
FHFA HPI$_{\{2006\}}$		-0.344***		0.176		0.247*
		(-3.21)		(1.52)		(1.79)
Mortgage Denial Rate$_{\{05\to06\}}$		-0.094		0.040		0.048
		(-1.51)		(0.35)		(0.42)
Mortgage Delinquency$_{\{05\to06\}}$		-0.050		-0.236***		-0.234***
		(-1.20)		(-4.69)		(-4.40)
#Counties	676	676	676	676	676	676
R-squared	0.344	0.452	0.124	0.365	0.122	0.356