NBER WORKING PAPER SERIES

WHEN REAL ESTATE IS THE ONLY GAME IN TOWN

Hyun-Soo Choi
Harrison Hong
Jeffrey Kubik
Jeffrey P. Thompson

Working Paper 19798
http://www.nber.org/papers/w19798

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2014

Choi acknowledges support from the Sim Kee Boon Institute, Singapore Management University. We thank Tim Loughran, Chris Mayer, Wenlan Qian, Jeremy Stein, and seminar participants at Dartmouth, Notre Dame, the CICF 2013 conference, and the NBER behavioural finance meeting for helpful comments. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff, the Board of Governors, or the National Bureau of Economic Research.

When Real Estate is the Only Game in Town
Hyun-Soo Choi, Harrison Hong, Jeffrey Kubik, and Jeffrey P. Thompson
NBER Working Paper No. 19798
January 2014
JEL No. G02,G11,G12,R21,R3

ABSTRACT

Using data on household portfolios and mortgage originations, we find that households residing in a city with few publicly traded firms headquartered there are more likely to own an investment home nearby. Households in these areas are also less likely to own stocks. This only-game-in-town effect is more pronounced for households living in high credit quality areas, who can access financing to afford a second home. This effect also becomes pronounced for households living in low credit quality areas after 2002 when securitization made it easier for these households to buy second homes. Cities with few local stocks have in equilibrium higher price-to-rent ratios, making it more attractive to rent, and lower (primary residence) homeownership rates.

Hyun-Soo Choi
Singapore Management University
Singapore
hschoi@smu.edu.sg

Harrison Hong
Department of Economics
Princeton University
26 Prospect Avenue
Princeton, NJ 08540
and NBER
hhong@princeton.edu

Jeffrey Kubik
Maxwell School
Syracuse University
426 Eggers Hall
Syracuse, NY 13244
jdkubik@maxwell.syr.edu

Jeffrey P. Thompson
Federal Reserve Board
Washington, DC
jeffrey.p.thompson@frb.gov

1. Introduction

According to the National Association of Realtors' Home Buyers Survey of 2012, investment homes, defined as homes bought for investment as opposed to occupation by the owner, represented around 22% of the residential sales market or around one million homes annually between 2003 and 2011. Investment homes are distinct from vacation homes, where the owner lives part-time. They are bought to generate rental income and typically face higher interest rates and more stringent collateral requirements than either primary residences or vacation homes because banks view them as speculative investments akin to stocks. Households characterize these purchases in the same way that they describe their purchases of IBM or Microsoft.

Investment homes are a sizeable part of the real estate market; and as can be seen in Table 1, demand for these homes played a big role in the recent housing cycle of 2002-2007. During the peak real estate bubble year of 2005, investment homes rose from typically being around 17% the market for residential sales to 28%. Around 12% were vacation homes and the remaining 60% were primary residences. Studies of this recent housing cycle, including Chinco and Mayer (2012), Li and Gao (2013), and Haughwout, Lee, Tracy, and van der Klaauw (2011), argue that speculative investment home purchases contributed to the the dramatic price run-ups in certain areas like Las Vegas and Phoenix. These findings suggest that studying the determinants of investment home or speculative purchases might be important for understanding real estate price cycles. Yet, the focus of research in the last twenty years in real estate economics has been on owner-occupied homes or primary residences.[1]

As such, we undertake a study of the determinants of investment home purchases in which we view these purchases through the lens of investing as opposed to owner occupation. Our

[1]For example, hedonics such as location or other housing amenities have been used to explain the cross-sectional variation in prices of primary homes (Glaeser (2007)). Moreover, even work on volume dynamics in housing markets and the rent versus buy decision have typically taken the motive for home purchases as one of owner occupation (Case and Shiller (1989), Stein (1995) and Sinai and Souleles (2005)).

analysis is motivated by a large literature on household finance that has greatly informed us on how households actually make investment decisions. We build on one of the most robust facts about household investment behavior, which is that they are locally biased. Households do not diversify but rather hold concentrated positions in stocks headquartered within 60 miles of where they live (see, e.g., Grinblatt and Keloharju (2001), Huberman (2001)). This local bias can be driven by familiarity bias or other informational frictions, whereby investors have both a small radius with which they search for investment opportunities and feel most comfortable with investments they know first hand. This local bias was originally discovered and characterized by French and Poterba (1991) as the international home bias puzzle for the lack of international diversification by equity investors.

Our premise is that investment homes are the ultimate locally-biased or home-biased investment. Familiarity bias or informational friction stories that characterize stock investments are even more relevant for investment homes because the owners have the tangibility of the home nearby. Indeed, real estate is often cast as an investment akin to gold in terms of its feeling of safety for households in under-developed or volatile financial markets. More concretely, one can think of *some* households as deriving utility from ownership, whether it is their primary residence or an investment home which they rent out. This utility might come from the status of ownership. Or it might come from the joys of having something to do such as collecting rents, much like a hobby. This latter motivation then provides an additional reason for why the local bias of investment homes is even stronger than equities since proximity to the investment home cuts down on the transportation cost of being a landlord.

Our hypothesis is that stocks headquartered nearby compete for the households' attention and lead to less ownership of investment homes. Their hobby becomes trading stocks rather than renting out investment homes. Importantly, we view our hypothesis as not primarily one of diversification since households tend not to diversify but one of attention and effort in which households derive utility from owning a home and the extra utility they get from homeownership is less if they have better substitutes such as trading stocks in the way described by Barber and Odean (2008).

Using data from the same National Realtors' Survey, we first document that purchases of investment homes are, if anything, more locally biased than equity investments. Panel B of Table 1 reports the distribution of distances between the primary residence and the owner's investment home versus vacation home. The median distance of the investment home from the buyers' primary residence is 25 miles. In contrast, the median distance of vacation homes is 305 miles. Indeed, Chinco and Mayer (2012) in a study of the top 20 MSAs confirm that most second home purchases are taken by households whose primary residence is in the same MSA. We will use this fact in interpreting our regressions below.

We then use two sources of data to test our hypothesis. The first is from an internal version, accessible only by Federal Reserve Bank economists, of the Survey of Consumer Finances (SCF), which samples a cross-section of roughly 5,000 to 6,000 households once every few years. We use the 1995, 1998, 2001, 2004, 2007, and 2010 waves. We know the MSA, zip-code and county where the household lives and have detailed data to construct their portfolio holdings including investment homes. We also have a host of demographic information to use in our analysis. We can distinguish between investment versus vacation homes and primary residences as well as capture the stock investments of these households. We calculate for each household whether or not they HAVE INVESTMENT REAL ESTATE, the value of its investment homes or vacation homes as a fraction of its total assets as well as its investments in stocks as a fraction of total assets (% INVESTMENT REAL ESTATE IN TOTAL ASSETS, % VACATION HOME IN TOTAL ASSETS, % DIRECTLY-HELD STOCKS IN TOTAL ASSETS).

The second is the Home Mortgage Disclosure Act (HMDA) mortgage origination data. HMDA keeps track of all mortgages originated in the U.S. from the period of 1998-2011. We measure as our dependent variable, %NON-OCCUPIED MORTGAGES, the amount of non-owner occupied home purchases as a fraction of total mortgages originated in a Metropolitan

Statistical Area (MSA) each year over this time period. While we do not know where the household that originated the mortgage lives, the survey evidence indicates that the preponderance of the investment homes are owned by households whose primary residence is also that MSA.

Our independent variable of interest, the supply of publicly available firms in an MSA, is the RATIO variable first used by Hong, Kubik, and Stein (2008). It is the ratio of the total book market value of firms headquartered in an MSA to the income in that MSA. They show that RATIO in an area is inversely related to the price-to-book of companies in that area. Moreover, most of the variation in their RATIO variable comes from the book value of firms headquartered in an area. That is, areas with fewer companies have higher prices for their stocks due to an only-game-in-town effect. In particular, they find that RATIO is closely correlated with population density in the state but is unassociated with the economic growth prospects in area. Their focus is on RATIO at the state level.

In our empirical designs involving the SCF and HMDA data, given that the bias of investment homes is even more local than for equity investments, we focus our analysis at the MSA level and even at the finer county level. Using the Survey of Consumer Finances data, we regress the fraction of a household's portfolio in investment real estates on household characteristics and the RATIO in the MSA that the household resides. We find effects that are economically important and statistically significant. Households in low RATIO cities have an investment real estate portfolio weight that is 15% higher than the mean of the sample. They also have a lower exposure to stocks than other households, consistent with our premise that households indeed look to real estate investments when there are few public equity opportunities around.

Of course, these regressions suffer from a potential omitted variables bias. To address this concern, we observe that our causal mechanism is premised on households being able to afford investment homes, which are obviously much more costly than simply opening a brokerage account and buying stocks. So if our relationship between public equity supply nearby and investment homes is really due to this substitution effect between investment homes and stocks, then we expect our only-game-in-town effect to be stronger for households with access to financing.

While we do not have data on the credit scores of our households, we build on earlier work by Mian and Sufi (2009), who find that the credit quality of households are clustered by MSA and county. We measure the credit quality of different MSAs and counties using the volume of sub-prime loan originations (identified using a standard methodology in the literature) in a county from the HMDA data. We can use this geographic information to identify the credit quality of households by the MSA or county where they live. We then calculate a difference-in-differences estimate in which we show that this only-game-in-town effect is more pronounced for households living in low sub-prime MSAs or counties where households can more easily access credit to afford investment homes to begin with.

Moreover, Mian and Sufi (2009) argue that certain sub-prime areas which did not get access to financing pre-2002 got significant access after 2002 as a result of securitization and the disintermediation of the credit process. Consistent with this perspective, Haughwout, Lee, Tracy, and van der Klaauw (2011) find that lax credit screening after 2002 allowed many households living in the sub-prime areas like Las Vegas and Phoenix to use a sub-prime loan to purchase second homes.

So we can then calculate a triple difference estimate in which we take advantage of the fact that there was a sub-prime lending boom after 2002, which Mian and Sufi (2009) and many others took to be a global shock that plausibly is exogenous to our relationship of interest. In other words, previously high sub-prime MSAs or counties started looking more like low sub-prime MSAs or counties after 2002. We show that high sub-prime MSAs or counties exhibited a significantly bigger only-game-in-town effect after 2002, consistent with our conjecture.

While the SCF data has detailed demographic information about households, we would like to know if this effect is important at an aggregate level. This is where the HMDA data is helpful since it contains most of the mortgage originations in the real estate market, especially in larger city centers or MSAs, which is the focus of our paper. Using the HMDA data, we regress the log of %NON-OCCUPIED MORTGAGES on the log of RATIO at the MSA level controlling for a host of MSA level characteristics such as unemployment and the elasticity of land supply as measured by Saiz (2010). We find that when the supply or RATIO is low, ownership of investment homes, %NON-OCCUPIED MORTGAGES, is high. We show graphically that the log specification fits the relationship between the ownership for investment homes in an MSA and the supply of stocks in that MSA very well as the relationship is highly non-linear.

A one standard deviation increase in the log of RATIO leads to a decrease in investment home demand that is 30% of the left-hand side variable's standard deviation. This economic effect is robust across different empirical specifications including one where we collapse our observations by averaging across the sample to obtain an average log %NON-OCCUPIED MORTGAGES and an average log RATIO and simply run a cross-sectional regression rather than a panel regression with clustering of standard errors by MSA. We then consider a number of robustness checks and confirm our results, including dealing with potential measurement error associated with mistaking investment for vacation homes or controlling for housing affordability in different areas. Moreover, we apply the same triple difference identification strategy and obtain similar results to those using SCF.

Given these HMDA results suggest that the only-game-in-town effect might have aggregate consequences, we develop a simple reduced model to understand the price implications of this only-game-in-town effect for investment homes. We assume that households have heterogeneous preferences for owning versus renting. Some prefer to own (perhaps for status or hobby reasons) while others prefer to rent even holding fixed prices for owning versus renting. Our only-game-in-town effect can be captured by the behavioral assumption that the utility households derived from owning is lower when there are more stocks nearby. With this reduced form behavioral assumption, we then solve for the equilibrium home prices and tenure choice in the presence of this only-game-in-town effect.

In equilibrium, when real estate is the only game in town, some households in an MSA are more apt to own an investment home (i.e. two homes including their primary residence) but that means some other households have to optimally want to rent rather than own. We show that this only-game-in-town demand for investment homes leads to higher price-to-rent ratios of homes in cities with few public firms, which then makes renting more attractive and allows markets to clear. Cities with few local stocks also then have lower (primary residence) homeownership rates, defined as the fraction of dwellings in a city in which the occupant is also the owner.

Importantly, we regress the log of the price-to-rent ratio on log RATIO and find a substantially higher price-to-rent ratios in MSAs with few local stocks. Moreover, consistent with our triple difference identification strategy above, we also find that high sub-prime cities with few local stocks experience the biggest price run-ups after 2002 when credit constraints got relaxed for them. We also gather data on homeownership rates from the Integrated Public Use Microdata Series (IPUMS) from 2005 to 2012. We regress the log of (primary) homeownership rate on log RATIO and find a substantially lower homeownership rates in MSAs with few local stocks, consistent with our hypothesis.

Our paper is related to a recent set of papers, Chinco and Mayer (2012), Li and Gao (2013), and Haughwout, Lee, Tracy, and van der Klaauw (2011), who collectively offer convincing evidence that households who buy investment homes behave more like speculators than households who buy for occupancy. Our study complements theirs in linking the second home purchase decision to local bias. Notably, Chinco and Mayer (2012) examine the impact of out of town buyers of second homes on home prices in 21 MSAs. Their study, consistent with the survey evidence and the premise of our study, documents that out of town buyers are typically a small fraction of second home purchases though it did rise significantly in some of the Sand or vacation cities like Phoenix and Las Vegas during the bubble period. This is one reason we drop the vacation MSAs in our HMDA analysis. Yet this small fraction of buyers was very informative for higher home price appreciation, which they attributed to out of town buyers being noise traders and their trend chasing causing housing bubbles.

Our paper is related to recent work on the speculative motive behind home purchases and home improvements including Choi, Hong, and Scheinkman (2013) and Gyourko and Saiz (2004). Glaeser, Gyourko, and Saiz (2008) also argue that the speculative motive behind the housing bubble was especially powerful in low supply elasticity states (such as New York and California where land is limited and zoning and development rules are stricter) in which supply could not quickly adjust to the rising home prices.

Our paper proceeds as follows. The results using the Survey of Consumer Finances data are presented in Section II. The empirical results using the HMDA mortgage origination data is presented in Section III. Equilibrium implications are modeled and tested in Section IV. We conclude in Section V.

2. Results Using Survey of Consumer Finances

2.1. Data

We use an internal version of the Survey of Consumer Finance which tells us at which MSA a household lives. As a survey of household finances and wealth, the SCF includes some assets that are broadly shared across the population (bank savings accounts) as well some that are held more narrowly and that are concentrated in the tails of the distribution (direct ownership of bonds). To support estimates of a variety of financial characteristics as well as the overall distribution of wealth, the survey employs a dual-frame sample design.

A national area-probability (AP) sample provides good coverage of widely spread characteristics. The AP sample selects household units with equal probability from primary sampling units that are selected through a multistage selection procedure, which includes stratification by a variety of characteristics, and selection proportional to their population.

Because of the concentration of assets and non-random survey response by wealth, the

SCF also employs a list sample which is developed from statistical records derived from tax returns under an agreement with Statistics of Income (SOI).[2] (See Kennickell (2000) for additional details on the SCF list sample.) This list sample consists of households with a high probability of having high net worth.[3]

The SCF joins the observations from the AP and list sample through weighting.[4] The weighting design adjusts each sample separately using all the useful information that can be brought to bear in creating post-strata. The final weights are adjusted so that the combined sample is nationally representative of the population and assets. These weights are used in all regressions.

2.2. Merging RATIO with SCF

Starting in the 2001, the public-use version of the SCF does not included any geographic identifiers. Prior to 2001 the public-use version only included the very broad 9-level Census Division code. The internal version of the SCF, however, has MSA-level (and lower levels of geography including zip-code and county) identification of where the household resides. RATIO is available at the MSA-level and is merged into the internal SCF data in the following way. A new MSA variable is created in the SCF using a state and county fips to MSA correspondence consistent with the geography used in calculating RATIO. The RATIO variable, for the years 1995, 1998, 2001, 2004, 2007, and 2010, is merged into the internal SCF data by MSA and year.

For the available years (triennial between 1995 and 2010), there are 363 total MSAs. Of those 363 MSAs, 133 are also included at some point in the SCF sample. 230 of those MSAs are not in the SCF sample. Not all households sampled in the SCF, however, reside in MSAs.

[2]See Wilson and William J. Smith (1983) and Internal Revenue Service (1992) for a description of the SOI file. The file used for each survey largely contains data from tax returns filed for the tax year two years before the year the survey takes place. See Kennickell (1998) for a detailed description of the selection of the 1998 list sample.

[3]For reasons related to cost control on the survey, the geographic distribution of the list sample is constrained to that of the area-probability sample.

[4]The evolution of the SCF weighting design is summarized in Kennickell (2000), with additional background by Kennickell and Woodburn (1992).

15 percent of SCF households are located in either rural areas, or in cities that are not part of MSAs. Of the MSAs that are in the SCF sample, 38 do not have RATIO data available.

Ultimately, there are 22,178 households in MSAs with RATIO data available sampled by the SCF between 1995 and 2010 that are used in this portion of the analysis. Because of the multiple imputation process (five implicates to generate a distribution for the imputed values) for missing values, there are 110,890 household-level records in the data. Standard errors in the SCF regressions are based on weighted data, and also are adjusted for the multiple implicates.

The key independent variable of interest is log RATIO, which is the log of the ratio of the total book value of firms headquartered in a MSA to the income in that MSA as in Hong, Kubik, and Stein (2008). Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take logs. The scaling is arbitrary and does not materially affect our conclusions. In additional analysis below, we experiment with different cut-offs on income. MSAs with higher income all have non-zero RATIO and our effects are if anything even stronger. As we show below, there is much more variation in the BOOK PER CAPITA across MSAs than there is in the INCOME PER CAPITA. This is consistent with the point in Hong, Kubik, and Stein (2008) that RATIO is essentially variation in BOOK PER CAPITA. We could have simply used this as the proxy for supply of stocks in each MSA instead of RATIO and simply control for INCOME PER CAPITA on the right hand side in making our inferences.

2.3. Definitions of Investment Real Estate and Vacation Homes

All dependent variables are represented as shares—they sum the dollar value of the households portion of all investment or vacation properties, and divide by either total assets or by total real estate assets. Investment real estates are aggregated from the following subcategories given by the SCF in the "property type" variable (x1703, x1803, and x1903): code 11 (land only: lot, tract, acreage; building lots, "farmland"); code 13 (substantial land

and other type of structure); code 15 (recreational property; sports field; golf course); code 24 (mobile home park); code 40 (one single family home); code 41 (multiple single family homes); code 42 (duplex 2 unit residence); code 43 (triplex); code 44 (fourplex); code 45 (5 or more); code 46 (apartment house, units unknown, rental units, or property nfs); code 47 (other business commercial property); code 48 (business/commercial and residential combination); code 49 (condo; co-op); and code 50 (residential, nec.).

%INVESTMENT REAL ESTATE IN TOTAL ASSETS is then the dollar value of investment real estates divided by the household's total assets. We also calculate %INVESTMENT REAL ESTATE IN REAL ESTATE ASSETS in which the denominator is the value of only the household's real estate assets as opposed to all its assets. Vacation homes are aggregated from the following categories reported by SCF: code 21 (seasonal/vacation); code 25 (time share); code 12 (substantial land and seasonal or other residence); and code 999 (other vacation home mapped from the mop-up). Analogously, we calculate %VACATION HOME IN TOTAL ASSETS. We also have data on how much households own of stocks. As such, we calculate % DIRECTLY-HELD STOCK HOLDINGS IN TOTAL ASSETS as the fraction of directly-held stock holdings in household's total asset.

We report in Table 2 the summary statistics of the Survey of Consumer Finance for 1995, 1998, 2001, 2004, 2007, 2010 waves. % INVESTMENT REAL ESTATE IN TOTAL ASSETS, the fraction of investment real estate in household's total asset, has a mean of 0.03 with a standard deviation of 0.09. % INVESTMENT REAL ESTATE IN REAL ESTATE ASSETS, the fraction of investment real estate in household's total real estate asset, has a mean of 0.08 with a standard deviation of 0.2. % VACATION HOME IN TOTAL ASSETS, the fraction of vacation home in household's total asset, has a mean of 0.01 with a standard deviation of 0.04. Among our sample of households, they are more apt to have an investment real estate than a vacation home. % DIRECTLY-HELD STOCKS IN TOTAL ASSETS, the fraction of directly-held stock holdings in household's total asset, has a mean of 0.02 with a standard deviation of 0.08.

In addition to these share variables, we also create related dummy variables to capture whether households have any investment real estate, vacation home or directly-held stocks. HAVE INVESTMENT REAL ESTATE equals one if the household has any investment real estate and zero other wise. The mean is 0.13 with a standard deviation of 0.34. So 13% of our households own some investment real estate. HAVE VACATION HOME equals one if the household has a vacation home and zero other wise. The mean is 0.06 with a standard deviation of 0.23. HAVE DIRECTLY-HELD STOCKS equals one if the household has any directly-held stock and zero otherwise. The mean is 0.21 with a standard deviation of 0.4. Overall, our households are more likely to have stocks than investment homes and more likely to have investment homes than vacation homes.

In Table 2, we also report summary statistics for our right-hand side variables. Log RATIO is the log ratio of the total book market value of firms headquartered in a MSA to the income in that MSA. The mean of Log RATIO is -1.53 with a standard deviation of 1.29. These figures are fairly similar to those reported for the HMDA sample below. This is reassuring since we ideally want comparable samples so that we can more easily evaluate the consistency of our results across HMDA and SCF. Even though we are measuring very different objects on the left hand side (mortgage originations versus household portfolios), we expect these numbers to be line up. UNEMP is unemployment rate from the Bureau of Labor Statistics. The mean is 5.54 and standard deviation is 2.25. HOME PRICE INDEX is the Federal Housing Finance Agency(FHFA) Housing Price Index at the MSA level has a mean of 160 with a standard deviation of 52.09. PAST PRICE APPRECIATION is the past home price appreciation over a year using the FHFA House Price Index, where $\frac{HPI_t - HPI_{t-1}}{HPI_{t-1}}$. It has a mean of 0.04 and a standard deviation of 0.06.

FAMILY SIZE is the number of people in each Primary Economic Unit. Our households have around 2.43 members with a standard deviation of 1.4. Log HOUSEHOLD INCOME has a mean of 10.84, which is around 51 thousand dollars. The SCF Survey is ideal for our study since we have higher net worth households who are most likely to be able to purchase an investment real estate. Table 7 also reports the breakdown of AGE of the head of each household. Most of our sample is below 55 years of age. But the sample is meant to be nationally representative and as such we see a distribution across the age cohorts. We also break down RACE. 73% of the households are white. 14% are black. Hispanics account for 9% of the households and Asians or others account for the remaining 4%.

In Table 2, we also break down EDUCATION. 13% have less than a high school education. 29% have high school or GED equivalent. 25% went to some college but do not have a Bachelor's degree. 20% have a Bachelor's degree. 7% have a Master's degree and 6% have a PhD. FAMILY STRUCTURE is reported. This breaks down situations like whether or not the couple is married and have children. Here LWP stands for the "living with partner" in the FAMILY STRUCTURE variable. We also have data on the NET WORTH of households in our sample, which we break up into quintiles or CATEGORY. We report the sum statistics by the five groups to give a sense of the range of the distribution. There is clearly significant variation in household net worth. Our richest CATEGORY 5 has a mean of around 3.5 million dollars with a standard deviation of around 10 million dollars. UNATTACHED FEMALE is the gender variable for Primary Economic Unit head, which is equal to 1 for un-partnered females and zero for everybody else. Around 27% of the households fit this designation with a standard deviation of 44%.

2.4. Portfolio Tilt to Investment Real Estate

We begin in Table 3 by reporting the household-level panel Logit regression results of our various dummy variables for investment real estate, vacation home and stock ownership. Household level control variables include AGE, EDUCATION, FAMILY STRUCTURE, RACE, UNATTACHED FEMALE, log HOUSEHOLD INCOME and NET WORTH CATEGORY. MSA-level control variables include unemployment rate, home price index, and past home price appreciation. AGE, EDUCATION, FAMILY STRUCTURE, and RACE are not reported for brevity.

In column (1), the dependent variable is the dummy variable HAVE INVESTMENT REAL ESTATE. The coefficient in front of log RATIO is -0.092 with a t-statistic of -2.52. Consistent with out only-game-in-town hypothesis, households are much less likely to own an investment real estate if they live in a high RATIO MSA. The marginal probability is -0.0087. Given that the unconditional probability is around 0.13, this is an economic significant effect. Continuing on Column (1), UNATTACHED FEMALE is also less likely to have an investment real estate. Higher household income increases the chances of having an investment real estate and so does having higher NET WORTH. The excluded group is the lowest NET WORTH CATEGORY 1. Notice that we can compare the economic significance of our log RATIO variable to these other covariates by comparing their marginal probabilities. For instance, household income has a marginal probability of around 0.024. So our marginal probability from log RATIO is around half the size of household income. Given that we would expect household income to be a significant determinant of having an investment real estate, our effect would seem to be economic interesting in comparison.

In column (2), we report the results for the propensity of households to have a vacation home. Notice now that log RATIO has no effect. But the other covariates such as log Household Income continue to have an effect. This is then reassuring that our RATIO variable is really picking up what we want in terms of influencing investment real estates.

In column (3), we report the results for the propensity to have directly-held stocks. Recall that our hypothesis postulates that in MSAs with fewer stocks nearby, households are less likely to have directly-held stocks in their portfolio. This is the premise of the substituting toward investment real estates in MSAs with few stocks. Consistent with our hypothesis, the coefficient of interest is 0.0776 with a t-statistic of 3.452. The marginal probability is around 0.01. We find indeed then that in MSAs with more stocks, households are more likely to own stocks. Since we already established in these MSAs that they are less likely to have investment real estate, we have now fully confirmed our hypothesis. Notice again that households with higher incomes are more likely to tilt toward stocks, while unattached females are less likely to tilt toward stocks.

One might worry that our findings in column (3) is somehow hard-wired. If we know that an MSA has fewer stocks, does not that mean that households there are less likely to be tilted toward stocks? So is there any information here? The investment real estate regressions result would seem to be more kosher in comparison. Notice that it is possible for the stocks headquartered in an MSA to be held by households in other areas, even if there is local bias. The effect of local bias is not one of complete segmentation. In other words, it is not a fait accompli that we had to find the result in column (3). Moreover, it is not obvious that with a lower supply of equities, the shares of these stocks could not be entirely held by locals, giving us a negative rather than a positive coefficient of interest.

In Table 4, we report the household-level panel Tobit regression results of % share of investment real estate, vacation home, and stock holdings in households' asset on log RATIO. Column (1) reports the regression result of % share of investment real estate in total asset(% INVESTMENT REAL ESTATE IN TOTAL ASSET) on log RATIO with household-level and MSA-level control variables.

Notice that the coefficient of interest in front of log RATIO is -0.0107 with a t-statistic of -1.652. This is a sizeable economic effect when we consider that we have in our survey very wealthy households with lots of assets. This effect lines up well with the fact that there are far fewer investment home mortgages originated in MSAs with lots of local equity investment opportunities. Some of the other covariates also attract statistically significant coefficients as before. Households with higher net worth also are more likely to have investment real estates. Unattached females are less likely to have investment real estate.

Column (2) reports the regression result of % share of investment real estate in total real estate asset (% INVESTMENT REAL ESTATE IN REAL ESTATE ASSET) on log RATIO and the same set of controls. Qualitatively, the results are similar to column (1). The coefficient of interest is -0.0167 with a t-statistic of 1.249. The significance is smaller than when we scale by total assets but the conclusion is similar. Again, the higher household net worth leads to greater assets held in investment real estates while unattached female leads to less. Comparing the economic significance across these key variables, we find that our only game in town effect is again sizeable when measured relative to household new worth.

Column (3) reports the regression result of % share of vacation home in total asset(% VACATION HOME IN TOTAL ASSET) on log RATIO with controls. Consistent with our earlier findings, we find no effect for vacation homes as we expect. Importantly, we now find that the higher is the UNEMP in an MSA the lower the vacation home purchases of households living there. Higher household new worth again contributes significantly to households owning a vacation home. One take-away from our analysis is that our only-game-in-town effect is really tied to investment real estates, especially if we contrast the fact that we still get a sizeable effect from household net worth on vacation home purchases.

Column (4) reports the regression result of % stocks holdings in total asset(%STOCK HOLDINGS IN TOTAL ASSET) on log RATIO and the same controls as in the earlier columns. We draw similar conclusions as in Table 11, confirming the ancillary prediction of our hypothesis that households when they own less investment real estates are more likely to own stocks.

2.5. Diff-in-Diff-in-Diff Estimates Using Credit Constraints

The premise of our hypothesis is that locally-biased households invest in local homes when stocks are not locally available. But of course, buying a home is much more expensive to simply opening a brokerage account and investing in some stocks. Credit constraints is a key factor, which is much more likely to affect households with lower income. So we can use credit constraints as a means to achieve better identification. We expect our only-game-in-town effect to be more prevalent for households living in high credit quality counties than in low credit quality counties. In other words, if we ran our earlier regressions separately for high and low credit quality MSAs or counties, we expect the bulk of our effect to be concentrated in high quality MSAs or counties.

We define the credit quality of an MSA or county using the volume of sub-prime lending from the HMDA data. We know from the HMDA data the loans that are labeled sub-prime. There are two major ways to identify sub-prime loans in the HMDA: 1) using the rate spread data in HMDA, or 2) using the Sub-prime and Manufactured Home Lenders list by the Department of Housing and Urban Development (HUD). Since the rate spreads are only available from 2004 and the HUD list is only available up to 2005, we identify the subprime by the HUD list before 2004 and use the rate spreads after 2004. As a result, we can calculate for different MSAs or counties the fraction of originations that are sub-prime in each year and then calculate the average sub-prime fraction for each MSA or county over the 2002-2007. Based on this average, we can create tercile groups based to label MSAs or counties as low, medium and high sub-prime MSAs or counties.

An indicator variable for a high sub-prime lending MSA or county is our measure of poor credit quality. Even better than this, we are fortunate that over our sample period the United States experienced a sub-prime credit lending boom for low income households. We know from the work of Mian and Sufi (2009) that sub-prime lending in certain MSAs or counties peaked in the years of 2002-2007. So we can do a diff-in-diff-in-diff estimate in which we expect our only-game-in-town effect to be bigger for high sub-prime county households in the years after 2002 than before 2002.

It is to an estimate of this result that we now turn in Table 5, where we formally estimate the triple-difference estimate. We create a variable I_Subprime MSA as being equal to 1 if the MSA is in the top tercile of MSAs in terms of sub-prime origination. We also have an indicator variable for after 2002, I_after2002 equals 1 if the year is after 2002 and zero otherwise. We then run our baseline regressions from Table 3 with HAVE INVESTMENT REAL ESTATE and % INVESTMENT REAL ESTATE IN TOTAL ASSETS on the left hand side and log RATIO, I_Subprime MSA, and I_after2002 and these three variables interacted. The coefficient of interest is in front of the double interaction term log RATIO * I_Subprime MSA and the triple interaction term log RATIO * I_Subprime MSA * I_after2002. We expect the coefficient in front of the double interaction term to be positive and the latter to be negative. These results are presented in Panel B. Instead of a subprime indicator, we also present in Panel A the same regression results where we sort MSAs into high, medium and low income based on the same tercile cut-offs for income. Subprime areas are fairly correlated with low income but they are not the same thing. We want to make this point clear in Panel A. We typically find stronger results when we use the subprime indicator than income indicator.

This is indeed what we find. From Panel B column (1) in which the dependent variable is HAVE INVESTMENT REAL ESTATE, notice that the coefficient of log RATIO is -.0127 with a t-statistic of 2.424. This is consistent with the earlier table in that households in MSAs with lots of supply of equities nearby are less likely to own investment real estate. The coefficient in front of the log RATIO * I_Subprime MSA term is positive with a coefficient of .0125 and a t-statistic of 1.395. This means that the effect of log RATIO is weaker in high subprime MSAs where households are less able to borrow, consistent with the premise of our identification strategy. This term is the right sign but not precisely estimated. This is somewhat expected in triple difference estimation strategies. The crucial term for us is the log RATIO * I_subprime MSA * I_after2002 term. This term attracts a negative coefficient of -.0228 with a t-statistic of -3.188. The economic and statical significance is large. Notice after 2002, high subprime MSA households experienced almost as big if not bigger only-game-in-town effect as low subprime MSA households.

We obtain the same results in column (2) when we use % of INVESTMENT REAL ESTATE IN TOTAL ASSETS. The coefficient in front of log RATIO is -.00259 with a t-statistic of -2.447. The coefficient in front of log RATIO * I_Subprime MSA is .00238 with a t-statistic of 1.53. Finally, the coefficient in front of the triple interaction term log RATIO * I_subprime MSA * I_after 2002 term is again negative at -.00445 with a t-statistic of 2.181. This coefficient is less precisely estimated than the one in column (1) but together they paint a supportive picture for our hypothesis. From Panel A, we see that we get largely similar results as Panel B, though less statistically and economically significant since income is not a perfect proxy for access to financing.

This set of diff-in-diff-in-diff estimates is consistent with our conjecture that the only-game-in-town effect is most pronounced for high credit quality households since lower credit quality households cannot borrow to buy homes even if there are no stocks headquartered in the neighborhood. This result speaks directly to our conjecture that the local supply of equity is affecting the demand for local investment homes. We can now interpret our OLS regressions more comfortably as causal since whatever omitted variables bias has to work through only high credit quality MSAs but not after 2002. In other words, the fact that MSAs with firms headquartered there also have less investment homes cannot be driven by omitted variables bias unless that bias also explains why it only affects high credit quality households which are the ones most likely to be able to borrow to buy investment homes but not after 2002 when even low credit quality households can borrow.

In Table 6, we refine our analysis in Table 5 even further by focusing on a finer unit of analysis, that of the county instead of an MSA. The HMDA data allows us to calculate volume of subprime originations at the county level. Using the same tercile splits as before, we find very similar results as using MSA. We conduct this analysis because we might worry that MSA is too broad a category for capture credit score features and credit quality of households might be more accurately captured at the county level.

2.6. Alternative Definitions of What Constitutes Investment Real Estate

As a robustness check of our results, we consider different definitions of investment real estate. First, we define investment real estate as strictly residential. Specifically, we consider the following subset of codes: code 24 (mobile home park); code 40 (one single family home); code 41 (multiple single family homes); code 42 (duplex 2 unit residence); code 43 (triplex); code 44 (fourplex); code 45 (5 or more); code 46 (apartment house, units unknown, rental units, or property nfs); code 48 (business/commercial and residential combination); code 49 (condo; co-op); and code 50 (residential, nec.). Second, the SCF asks whether the household or Primary Economic Unit (PEU) earns any money (and how much) from other real estate. This definition of investment real estate simply takes the asset value for those other properties that are generating income for the PEU. We find that both definitions show similar results to our original definition above.

3. Results Using HMDA

3.1. Data

Our data on mortgage originations comes from the Home Mortgage Disclosure Act (HMDA). This data covers the period of 1998 to 2011 and includes all mortgage originations in the U.S. by MSA. This data further breaks down whether the mortgage is intended for a primary residence or a non-owner occupied house. It also gives us detailed information on the mortgage contract such as the interest rate and loan value. We also use the COMPUSTAT for the total book value of firms headquartered in a city. We use the home price index from the Federal Housing Finance Agency and the fair market rent from the U.S. Department of Housing and Urban Development. Our other demographic data are from the Bureau of Labor Statistics and the Bureau of Economic Analysis.

Table 7 collects the summary statistics for our main variables of interest. The summary statistics are reported by MSA level. The NON-OWNER OCCUPIED HOME PURCHASE AMTO has a mean of 198 million dollars with a standard deviation of 603 million dollars. The TOTAL HOME PURCHASE AMTO for all mortgage originations for home purchase regardless of type is 1.67 billion dollars with a standard deviation of 5.24 billion dollars. The %NON-OCCUPIED MORTGAGES is the ratio of the non-owner occupied originations for home purchase to total originations for home purchase in each MSA. It has a mean of 14% with a standard deviation of 10%. This is our dependent variable of interest.

The mean of log RATIO is -6.33 with a standard deviation of 4.68. We also report the

BOOK PER CAPITA, which is the total book value of firms headquartered in each MSA divided by the population of that MSA. BOOK PER CAPITA has a mean of 5355 dollars with a standard deviation of 16,474 dollars. INCOME PER CAPITA has a mean of 31,463 dollars with a standard deviation of 7222 dollars. In addition, we report summary statistics for the key covariates. PAST PRICE APPRECIATION has a mean of 3.18% per annum with a standard deviation of 6.65%. The log of population, log POP, has a mean of 12.63 with a standard deviation of 1.06. Unemployment, UNEMP, has a mean of 5.92% and a standard deviation of 2.78%. Elasticity is the Saiz (2010) land supply elasticity index, which captures how easy it is to build. The higher the elasticity the easier it is to build. It has a mean of 2.52 with a standard deviation of 1.42. The Housing Affordability Index is from the National Association of Realtors, which is available only from 2009 to 2011.[5] The Housing Affordability Index measures whether or not a family with median income could qualify for a mortgage loan on a median-priced home. Higher index indicates that a median-priced home is more affordable to a median income family. The Housing Affordability Index has a mean of 130.6 with a standard deviation of 38.83. Finally, the log of the price-to-rent ratio, log PR, has a mean of -1.5 with a standard deviation of 0.23.

3.2. Baseline Regressions

In Table 8, we report the baseline regression results of the log of the percentage of non-owner occupied mortgage originations (log %NON-OCCUPIED MORTGAGES) on log RATIO. Columns (1)-(3) report the panel regression estimates using MSA-level samples with yearly observations from 1998 to 2011. Column (1) reports the univariate regression. The coefficient on RATIO is -0.0366 with a t-statistic of -7.481. The standard errors are clustered by MSA. The economic significance is calculated as a one-standard deviation movement of the log RATIO times the coefficient of interest divided by the standard deviation of the left hand side variable. In column (1), a one standard deviation increase in log RATIO leads

[5] http://www.realtor.org/topics/housing-affordability-index

to a fall of log %NON-OCCUPIED MORTGAGES that is 28% of its standard deviation. This is a sizeable move. Columns (2)-(3) report the multivariate regression results. We introduce controls including year fixed effects, unemployment rate (UNEMP) and housing supply elasticity (Elasticity) from Saiz (2010). These specifications produce almost identical estimates to the univariate regression. It is interesting to observe from column (3) that MSAs with lower supply elasticity have higher %NON-OCCUPIED MORTGAGES. But this covariate only makes our only-game-in-town effect stronger since our coefficient rises to -0.04 with a t-statistic of -5.9 in column (3).

In columns (4)-(6), rather than running a panel regression, we take the average of all the variables over our sample period and report the results of a pure cross-sectional regression. Hong, Kubik, and Stein (2008) show that RATIO is highly persistent over time; hence, collapsing the panel regression in a pure cross-section might generate a better specification. Column (4) reports the univariate regression. The coefficient is -0.0603 with a t-statistic of -4.406. The economic significance is now -0.23. That is, a one standard deviation movement in log RATIO leads to a fall of log %NON-OCCUPIED MORTGAGES that is 23% of the standard deviation of the left-hand side variable. Columns (5)-(6) report the multivariate regression results and we obtain similar results to those in column (4).

In Figure 1, we plot the relationship between log %NON-OCCUPIED MORTGAGES on the y-axis and the log RATIO on the x-axis for the regression specification in column (4). One sees a very clear linear and downward sloping relationship between these two variables. Indeed, this implies that a plot of the levels of these two variables (raw rather than in logs) is non-linearly and negatively related. Moreover, notice that since we have many MSAs with zeros, there is a vertical line of observations to the far left of the graph. It is easy to see that these observations are not qualitatively driving our results since removing the vertical line would not significantly change the inference that the slope of the fitted line is negative. We show that this is the case below.

3.3. Decomposing RATIO

In Table 9, we report the regression results in which we decompose the RATIO variable into BOOK PER CAPITA and INCOME PER CAPITA. The left-hand side variable (log %NON-OCCUPIED MORTGAGES) is the log ratio of non-owner-occupied mortgage originations for home purchase to total mortgage origination for home purchase. Right-hand side variables are now the log of book value per capita (log BOOK PER CAPITA) and the log of per capita income (log INCOME PER CAPITA). In other words, rather than regressing log %NON-OCCUPIED MORTGAGES on log RATIO, we regress log %NON-OCCUPIED MORTGAGES instead on log BOOK PER CAPITA, the numerator of log RATIO, and log INCOME PER CAPITA, the denominator of log RATIO. This regression specification tells us whether the effects of RATIO on %NON-OCCUPIED MORTGAGES is coming from the supply of public equity, that is the BOOK PER CAPITA, or the income of the region, INCOME PER CAPITA.

Columns (1)-(2) report the panel regression using MSA-level samples from 1998 to 2011. Control variables include unemployment rate (UNEMP) and housing supply elasticity from Saiz (2010). The estimates in column (1) suggest that all the effects are coming from log BOOK PER CAPITA. The coefficient on log BOOK PER CAPITA in column (1) is -0.0129 with a t-statistic of -4.549. The implied economic significance is 21% of the left-hand side variable. In contrast the coefficient on log INCOME PER CAPITA is not statistically significant. In column (2), the coefficient of log INCOME PER CAPITA becomes statistically significant but is wrong-signed when we include Elasticity, however. We expect that MSAs with higher income have a lower RATIO (or less effective supply of stocks to go around given the income) and hence more investment home purchases. Instead, we find that high per capita income MSAs have less investment home purchases.

Columns (3)-(4) report the pure cross-sectional regression on the averaged observations across the sample period. Again, we see that the coefficient on log BOOK PER CAPITA is large and statistically significant. The coefficient on log INCOME PER CAPITA is again wrong-signed and insignificant. But the coefficient becomes statistically significant when we include Elasticity.

The results in this table show our results are not driven by strange specification assumptions. It is fairly robust to different parameterizations, with most of the effect coming from the supply of public equities as captured by log BOOK PER CAPITA.

3.4. Additional Controls for Population Density, Price Appreciation, and Housing Affordability Index

In Table 10, we augment our baseline regression specification from Table 8 by examining what happens when we control for a variety of factors. In column (1), we add population density of a city, measured by log of MSA population, to our regression specification. Hong, Kubik, and Stein (2008) show that log POP and RATIO are highly correlated because the book value of firms are lower in less populated areas. They argue that log POP might be an instrument for RATIO.

We report in column (1) the regression results of the percentage of non-owner occupied mortgage originations on RATIO, controlling the log number of population (logPOP). Column (1) reports the panel regression using MSA-level samples from 1998 to 2011. Controls include unemployment rate (UNEMP) and housing supply elasticity (Elasticity). The coefficient on log RATIO (-0.0275) is still statistically and economically significant even with log POP in the regression. The coefficient on log POP (-0.107) is also statistically significant and attracts the same sign as log RATIO as we expect given the findings in Hong, Kubik, and Stein (2008).

In column (2), we report the regression results of the percentage of non-owner occupied mortgage originations on RATIO after dropping vacation MSAs. Vacation MSAs are the popular vacation home areas listed from CNBC, Barron, and Forbes. Column (2) report the panel regression using MSA-level sample from 1998 to 2011. The coefficient of interest is -0.0391 with a t-statistic of -6.048. It is virtually identical to that reported in Table 8.

In columns (3)-(4), we report the regression results of the percentage of non-owner occupied mortgage originations on RATIO, controlling for home price appreciation. The idea is that the amount of investment home purchases are likely to depend and be positively correlated with home price movements. We want to see whether our coefficient on log RATIO is picking up some of the explanatory power of these trends in home prices. This is unlikely because RATIO is quite persistent over time, but it would nonetheless be comforting to verify.

In column (3), we control the future home price appreciation(FUTURE PRICE APPRECIATION), where FUTURE PRICE APPRECIATION$_t = \frac{HPI_{t+1} - HPI_t}{HPI_t}$ (where HPI is the annual house price index value of the MSA). Again, the coefficient of interest in front of log RATIO is if anything a bit stronger than the results presented in Table 8.

In column (4), we control for past home price appreciation(PAST PRICE APPRECIATION), where PAST PRICE APPRECIATION$_t = \frac{HPI_t - HPI_{t-1}}{HPI_{t-1}}$. We get statistically and economic significant estimates of the coefficient on log RATIO that are again almost identical to the results presented in Table 8.

In columns (5)-(8), we report the regression results of the percentage of non-owner-occupied mortgage originations on RATIO, controlling for the Housing Affordability Index from the National Association of Realtors. We want to see how the investment home purchases are related to the regional mortgage availability and how our coefficient on log RATIO is affected. We use the Metropolitan series of the Housing Affordability Index which gives us affordability in different MSAs.

In columns (5)-(6), we control for this index. Other control variables include unemployment rate(UNEMP) and housing supply elasticity(Elasticity). Again, the coefficient of interest in front of log RATIO is if anything a bit stronger than the results presented in Table 8. We also find that the coefficient on the Housing Affordability Index is negative and statistically significant. This means that the percentage of non-owner-occupied mortgage originations is negatively associated with housing affordability in the city. However, the coefficient becomes insignificant when we include Elasticity.

In columns (7)-(8), we take the time series average of this Metropolitan Index for years of 2009 to 2011 and assume that in earlier years each MSA has an affordability index equal to their 2009 to 2011 time series average. We call this Housing Affordability Index Backfilled. We find similar results.

3.5. Owner Occupied Originations

In Table 11, instead of the fraction of investment home mortgages to total mortgages on the left-hand side, we study instead the amount of owner-occupied mortgage originations. The idea is for us to check to see whether log RATIO influences owner-occupied originations at all. We ideally would not want to find any effect since our hypothesis pertains only to non-owner occupied homes. If we found a significant effect, this might raise other alternative stories.

Our prior is that RATIO should have no effect on primary residences since we know RATIO is inversely related to the fraction of investment home mortgages to total mortgage origination. Since total mortgage origination is the sum of owner-occupied and non-owner occupied mortgages, the fact that we find RATIO has an effect on this fraction should mean it should have little effect on owner occupied originations. Otherwise this movement might offset any effects we see in non-owner occupied originations.

But just to check, we report the regression results of the owner-occupied mortgage originations on RATIO in Table 11. The left-hand side variable (log OWNER OCCUPIED MORTGAGES) is the log amount of owner-occupied mortgage originations for home purchase. The independent variables are the same as before. Columns (1)-(3) report the panel regression using MSA-level samples from 1998 to 2011. Column (1) reports the univariate regression. The coefficient is 0.172 with a t-statistic of 14.39.

However, this coefficient declines significantly once we control for covariates in columns (2)-(3). The effect of log RATIO on owner-occupied originations as we expect is negligible.

While the effect is very precisely measured, the economic significance is tiny. The coefficient of interest in column (3) is 0.0106 and the economic significance is a tiny 0.03. Columns (4)-(6) report the cross-sectional regression by averaging observations across the sample period. The results largely reinforce the fact that RATIO primarily influences non-owner occupied mortgage origination.

3.6. Diff-in-Diff-in-Diff Estimates Using Credit Constraints

It is to a more formal estimate of this result that we now turn in Table 12, where we formally estimate the triple-difference estimate for Table 5. We create a variable I_Subprime MSA as being equal to 1 if the MSA is in the top tercile of MSAs in terms of sub-prime origination. We also have an indicator variable for after 2002, I_after2002 equals 1 if the year is after 2002 and zero otherwise. We then run our baseline regressions with log of non-owner occupied mortgages on the left hand side and log RATIO, I_Subprime MSA and I_after2002 and these three variables interacted. The coefficient of interest is in front of the triple interaction term: log RATIO * I_Subprime MSA * I_after2002. We expect the coefficient in front of the term to be negative. That is, we expect that log RATIO has more of a negative effect for high sub-prime MSAs compared to low sub-prime MSAs after 2002. This is indeed what we find. The coefficient in front of the term has a coefficient of -0.0131 and a t-statistic of -1.692. It is very comforting that both triple difference estimates of interest are economically and statistically significant, just as we had found earlier using the SCF data.

4. Pricing and Homeownership Rate Implications

Having established that households living in MSAs with less local stocks are more likely to buy investment homes, we turn to an analysis of the equilibrium implications of local bias of investment homes. This equilibrium analysis necessarily involves the modeling of tenure choice (renting versus owning) and investment home choice (owning one versus two homes).

Our model is a simple extension of Sinai and Souleles (2005).

We assume that there are two periods, $t = 0, 1$. There are N identical homes. Assume that it costs P_0 to own a house at $t = 0$ and it can be sold at $\tilde{P}_1$ at $t = 1$, with mean $E(\tilde{P}_1)$ and variance $Var(\tilde{P}_1)$. A house can be rented out at R from $t = 0$ to $t = 1$. There are N households, who have mean-variance expected utility with γ. Each household has to either own or rent one of the homes. If a home is rented, then one of the households must own two homes. Assume that there exist two types of households, in equal fractions in the economy: those who get extra utility from owning home (O) and those who does not get any extra utility from owning home (R). That is, O-type agents will get extra expected return from owning a home, U_O. The extra expected return is lower for second home, αU_O with $0 < \alpha < 1$.

All households are endowed with local stock wealth $(\tilde{Q})$, with mean $E(\tilde{Q}) = 0$ and variance $Var(\tilde{Q}) = Var(b\tilde{u}) = b^2\sigma^2$, where b is constant and $Var(\tilde{u}) = \sigma^2$. The parameter b corresponds to the supply of local stocks in an MSA. Small b corresponds to lower supply MSAs. We assume that stock values and home prices are positively correlated, $Cov(\tilde{P}_1, \tilde{Q}) > 0$.

Our only-game-in-town effect for investment homes, as motivated in the Introduction, corresponds to the following assumption: α, the utility obtained from the investment home is a decreasing function of b. When there are more stocks locally, the utility households derive from owning an investment home are lower.

We solve for the equilibrium home prices and tenure choice in the presence of this only-game-in-town effect. In equilibrium, when real estate is the only game in town, O-households in an MSA are more apt to own an investment home (i.e. two homes including their primary residence) but that means R-households have to optimally want to rent rather than own. We show that this only-game-in-town demand for investment homes leads to higher price-to-rent ratios of homes in cities with few public firms, which then makes renting more attractive and allows markets to clear. It also leads to a lower homeownership rate, defined as the fraction of dwellings or homes in a city in which the occupant is also the owner. Note that someone who owns a primary residence and an investment home is only counted as a homeowner for his primary residence but his investment home is non-owner occupied or counted as rented.

For the equilibrium with R-type agents renting and O-type agents owning 2 homes in equilibrium, the following three conditions should hold. First, Equation (1) below indicates the condition so that R-type agents prefer renting to owning. Second, Equation (2) indicates the condition such that O-type agents will prefer owning to renting. Third, O-type agents must be indifferent between owning two houses and owning one house, which is Equation (3).

$$E_0[U^R(W - R + \tilde{Q})] > E_0[U^R(W - P_0 + \tilde{P}_1 + \tilde{Q})] \tag{1}$$

$$E_0[U^O(W - P_0 + \tilde{P}_1 + \tilde{Q})] > E_0[U^O(W - R + \tilde{Q})] \tag{2}$$

$$E_0[U^O(W - 2P_0 + 2\tilde{P}_1 + R + \tilde{Q})] = E_0[U^O(W - P_0 + \tilde{P}_1 + \tilde{Q})] \tag{3}$$

To get the equilibrium price, we can solve Equation (3) to obtain the following:

$$W - 2P_0 + 2E(\tilde{P}_1) + R + (1 + \alpha(b))U_O - \frac{\gamma}{2}(4Var(\tilde{P}_1) + 4Cov(\tilde{P}_1, \tilde{Q}) + Var(\tilde{Q}))$$
$$= W - P_0 + E(\tilde{P}_1) + U_O - \frac{\gamma}{2}(Var(\tilde{P}_1) + 2Cov(\tilde{P}_1, \tilde{Q}) + Var(\tilde{Q}))$$

$$P_0 = R + E(\tilde{P}_1) + (\alpha(b)U_O - \gamma Var(\tilde{P}_1)) - \frac{\gamma}{2}Var(\tilde{P}_1) - \gamma Cov(\tilde{P}_1, \tilde{Q}) \tag{4}$$

Equation (2) will be satisfied if the equilibrium price satisfies the following condition.

$$W - P_0 + E(\tilde{P}_1) + U_H - \frac{\gamma}{2}(Var(\tilde{Q}) + 2Cov(\tilde{P}_1, \tilde{Q}) + Var(\tilde{P}_1)) > W - R - \frac{\gamma}{2}Var(\tilde{Q})$$
$$\rightarrow P_0 < R + E(\tilde{P}_1) + U_O - \frac{\gamma}{2}Var(\tilde{P}_1) - \gamma Cov(\tilde{P}_1, \tilde{Q})$$

We can show that the equilibrium price always satisfies the condition. That is, O-type agents will always prefer owning to renting at the price.

Using Equation (1), we get our main condition for the equilibrium price such that the R-agents will want to rent:

$$W - R - \frac{\gamma}{2}Var(\tilde{Q}) > W - P_0 + E(\tilde{P}_1) - \frac{\gamma}{2}(Var(\tilde{Q}) + 2Cov(\tilde{P}_1, \tilde{Q}) + Var(\tilde{P}_1))$$

$$\rightarrow P_0 > R + E(\tilde{P}_1) - \frac{\gamma}{2}Var(\tilde{P}_1) - \gamma Cov(\tilde{P}_1, \tilde{Q}) \tag{5}$$

Comparing Equation (4) and (5), we get

$$\alpha(b)U_O - \gamma Var(\tilde{P}_1) > 0 \tag{6}$$

as the condition for the R-type to rent.

As b increases, P_0 decreases in equation (4) since 1) O-type agents get less utility from owning second home, and 2) the covariance between stock value and housing price becomes higher. Equation (6) becomes harder to hold as b increases, that is, R-type agents are less likely to be a renter. When b is high enough to have $\alpha(b)U_O - \gamma Var(\tilde{P}_1) \leq 0$, R-type agents will start to own a house. It is easy to prove the following proposition which forms the crux of our remaining empirical work.

Proposition 1. *The price-to-rent ratio rises and the (primary residence) homeownership rate falls when there is less supply of local equity investment opportunities.*

If in fact investors suffer from local bias and invest in second homes nearby when there is few equity investment opportunities around, we expect these low RATIO places to have higher home prices. In Table 13, we see if this is true by regressing the price-to-rent ratio for homes in an MSA on the MSA's RATIO. The left-hand side variable (log PR) is the log of price to rent ratio. log RATIO is the log ratio of the total book market value of firms headquartered in a MSA to the income in that MSA.

Columns (1)-(3) report the panel regression using MSA-level samples from 1990 to 2011.

Column (1) reports the univariate regression. The coefficient of interest is -0.004 with a t-statistic of -2.709. The economic significance is -0.08 of the left-hand side variable's standard deviation. This is is a sizeable effect. In Figure 2, we make a plot of this regression to show the fitted line through the data. We see that there is a nice downward sloping relationship between log (PR) and log RATIO.

Columns (2)-(3) report the multivariate regression results. Control variables include unemployment rate (UNEMP) and housing supply elasticity (Elasticity). The coefficient of interest gets stronger and gets to as high as a -0.22 of economic significance. Columns (4)-(6) report the cross-sectional regression by averaging observations across the sample period. Column (4) reports the univariate regression. Columns (5)-(6) report the multivariate regression results. The effects are stronger as before using these specifications. The coefficient of interest from column (5) is -0.0179 with a t-statistic of -7.697. The economic significance is now -0.39.

Moreover, consistent with our triple difference identification strategy above, we also expect to find that high sub-prime cities with few local stocks ought to experience the biggest price run-ups after 2002 when credit constraints got relaxed for them. There is in this prediction an assumption that the easiness of credit which primarily affected housing also eased for refinancing homes which allowed for greater liquidity to engage in other types of investments and also consumption. From Figure 3, we can see that home price appreciation pre 2007 was greatest in high subprime areas. This is controlling for income as we see that high subprime areas with income terciles uniformly experienced greater home price appreciation than low subprime areas. In Table 14, we then regress home price appreciation during the periods of 1998-2005, 2002-2005 and 2002-2007, columns (1)-(3) respectively, on log RATIO, I_subprime and log RATIO*I_subprime. Notice that the coefficient in front of I_subprime is positive but the interaction term is negative. So most of the home price appreciation is occurring in low RATIO areas, that is areas where real estate is the only game in town.

From Proposition 1, when investors invest more in second homes nearby with few equity investment around, we also expect these low RATIO places to have lower (primary residence) homeownership rates. In Table 15, we see if this is true by regressing the homeownership rate in an MSA on the MSA's RATIO. The left-hand side variable (log HOMEOWNERSHIP RATE) is the log of homeownership rate. We construct MSA-level homeownership rate from the Integrated Public Use Microdata Series (IPUMS).[6] The IPUMS consists of more than fifty high-precision samples of the American population drawn from fifteen federal censuses and from the American Community Surveys of 2000-2012. The data include the ownership of dwelling, for about 2 millions household samples per year, whether the housing unit was rented or owned by its inhabitants. Using the MSA code of the household samples which is only available from 2005, we construct MSA-level (primary residence) homeownership rate from 2005 to 2011 for 266 MSAs.[7] Constructed homeownership rate variable shows a mean of 68.67% and a standard deviation of 6.26.

Panel A reports the pooled panel regression using MSA-level samples from 2005 to 2011. Control variables include unemployment rate (UNEMP), housing supply elasticity (Elasticity) and log number of population (logPOP). In column (2), the coefficient of interest is 0.0042 with a t-statistic of 2.65. The economic significance is 0.2 of the left-hand side variable's standard deviation, that is 0.095. This is a sizeable effect. Panel B reports the cross-sectional regression by averaging observations across the sample period and Panel C reports the Fama-MacBeth regression. The results are very similar to the pooled panel regression.

5. Conclusion

There is now great interest in understanding the sources of fluctuations in home prices around the world between the years of 2002 to 2007. Some argue that a toxic brew of lower interest rates, easy credit and agency problems might have contributed to the boom and

[6] See Ruggles, Alexander, Genadek, Goeken, Schroeder, and Sobek (2010)

[7] MSA-level homeownership rates are available from Census for 75 largest MSAs. Our measure of homeownership from the IPUMS is highly correlated to the Census for those 75 MSAs.

bust patterns we saw during this time (see, e.g., Himmelberg, Mayer, and Sinai (2005), Taylor (2009), Khandani, Lo, and Merton (2013), Mian and Sufi (2009), Keys, Mukherjee, Seru, and Vig (2009)). Others argue that behavioral triggers, whether they be optimistic beliefs (see Shiller (2005), Glaeser, Gottlieb, and Gyourko (2013), Greenspan (2010), Bubb and Kaufman (2009), Cheng, Raina, and Xiong (2013)) or momentum trading in investment homes (Chinco and Mayer (2012), Li and Gao (2013), and Haughwout, Lee, Tracy, and van der Klaauw (2011)) were more critical. Our paper suggests that an only-game-in-town-effect might have exacerbated these important forces. The housing bubble came with the end of the Internet Bubble and a period of little initial public offerings or equity supply. Whether this was a coincidence or the lack of equity supply might have exacerbated the boom-bust housing cycle of 2002-2007 is an interesting question for future research. More broadly, investments in equities and investments in real estate are treated as separate endeavors by researchers. But our analysis suggests that they are likely to be jointly determined and understanding their interactions would significantly increase our understanding of investors' portfolio choices and the impact of these choices on these critical financial markets.

References

Barber, B. M., and T. Odean, 2008, "All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors," *Review of Financial Studies*, 21(2), 785–818.

Bubb, R., and A. Kaufman, 2009, "Securitization and Moral Hazard: Evidence from Credit Score Cutoff Rules," *SSRN Working Paper*, pp. 1–73.

Case, K. E., and R. J. Shiller, 1989, "The Efficiency of the Market for Single-Family Homes," *American Economic Review*, 79(1), 125–137.

Cheng, I.-H., S. Raina, and W. Xiong, 2013, "Wall Street and the Housing Bubble," *NBER Working Paper*, pp. 1–39.

Chinco, A. M., and C. J. Mayer, 2012, "Distant Speculators and Asset Bubbles in the Housing Market," *NBER Working Paper*, pp. 1–52.

Choi, H.-S., H. Hong, and J. Scheinkman, 2013, "Speculating on Home Improvements," *Journal of Financial Economics*, , forthcoming.

French, K. R., and J. M. Poterba, 1991, "Investor Diversification and International Equity Markets," *American Economic Review*, 81(2), 222–226.

Glaeser, E. L., 2007, "The Economics Approach to Cities," *NBER Working Paper*, pp. 1–39.

Glaeser, E. L., J. D. Gottlieb, and J. Gyourko, 2013, "Can Cheap Credit Explain the Housing Boom?," *Housing and the Financial Crisis, University of Chicago Press*, pp. 301–359.

Glaeser, E. L., J. Gyourko, and A. Saiz, 2008, "Housing Supply and Housing Bubbles," *Journal of Urban Economics*, 64(2), 198–217.

Greenspan, A., 2010, "The Crisis," *Brookings Papers on Economic Activity*, pp. 201–246.

Grinblatt, M., and M. Keloharju, 2001, "How Distance, Language, and Culture Influence Stockholdings and Trades," *Journal of Finance*, 56(3), 1053–1073.

Gyourko, J., and A. Saiz, 2004, "Reinvestment in the Housing Stock: the Role of Construction Costs and the Supply Side," *Journal of Urban Economics*, 55(1), 238–256.

Haughwout, A., D. Lee, J. Tracy, and W. van der Klaauw, 2011, "Real Estate Investor, the Leverage Cycle, and the Housing Market Crisis," *Federal Reserve Bank of New York Staff Reports 514*.

Himmelberg, C., C. Mayer, and T. Sinai, 2005, "Assessing High House Prices: Bubbles, Fundamentals and Misperceptions," *Journal of Economic Perspectives*, 19(4), 67–92.

Hong, H., J. D. Kubik, and J. C. Stein, 2008, "The Only Game in Town: Stock-Price Consequences of Local Bias," *Journal of Financial Economics*, 90(1), 20–37.

Huberman, G., 2001, "Familiarity Breeds Investment," *Review of Financial Studies*, 14(3), 659–680.

Internal Revenue Service, 1992, *Individual Income Tax Returns, 1990*Internal Revenue Service.

Kennickell, A. B., 1998, "List Sample Design for the 1998 Survey of Consumer Finances," *FRB Working Paper*.

——— , 2000, "Wealth Measurement in the Survey of Consumer Finances: Methodology and Directions for Future Research," *FRB Working Paper*.

Kennickell, A. B., and R. L. Woodburn, 1992, "Estimation of Household Net Worth Using Model-Based and Design-Based Weights: Evidence from the 1989 Survey of Consumer Finances," *FRB Working Paper*.

Keys, B. J., T. Mukherjee, A. Seru, and V. Vig, 2009, "Financial Regulation and Securitization: Evidence from Subprime Mortgage Loans," *Journal of Monetary Economics*, 56(5), 700–720.

Khandani, A. E., A. W. Lo, and R. C. Merton, 2013, "Systematic Risk and the Refinancing Ratchet Effect," *Journal of Financial Economics*, 108, 29–45.

Li, W., and Z. Gao, 2013, "Real Estate Investors and the Boom and Bust of the US Housing Market," *SSRN Working Paper*, pp. 1–38.

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.

Ruggles, S., J. T. Alexander, K. Genadek, R. Goeken, M. B. Schroeder, and M. Sobek, 2010, "Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]," *Minneapolis: University of Minnesota, 2010.*

Saiz, A., 2010, "The Geographic Determinants of Housing Supply," *Quarterly Journal of Economics*, 125(3), 1253–1296.

Shiller, R. J., 2005, "Irrational Exuberance, 2nd Edition," *Princeton University Press.*

Sinai, T., and N. S. Souleles, 2005, "Owner-occupied Housing as a Hedge Against Rent Risk," *Quarterly Journal of Economics*, 120(2), 763–789.

Stein, J. C., 1995, "Prices and Trading Volume in the Housing Market: A Model with Down-Payment Effects," *Quarterly Journal of Economics*, 110(2), 379–406.

Taylor, J. B., 2009, "Getting Off Track: How Government Actions and Interventions Caused, Prolonged, and Worsened the Financial Crisis," *Hoover Institution Press.*

Wilson, O. H., and J. William J. Smith, 1983, "Access to Tax Records for Statistical Purposes," *Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 595–601.

Table 1: Investment and Vacation Home Purchase

Panel A reports the share of home sales by intended use from the *Investment and Vacation Home Buyers Survey* by the National Association of Realtors. Panel B reports the distance of second homes from primary residence from the *Investment and Vacation Home Buyers Survey*.

Panel A : *Share of Home Sales by Intended Use*

	Primary Residences (%)	Vacation Properties (%)	Investment Properties (%)
2003	67	12	22
2004	64	11	25
2005	60	12	28
2006	64	14	22
2007	67	12	21
2008	70	9	21
2009	73	10	17
2010	73	10	17
2011	61	11	27

Panel B : *Distance from Primary Residence*

	Primary Residences (%)	Vacation Properties (%)	Investment Properties (%)
5 miles or less	-	2	17
6 to 10 miles	-	3	9
11 to 15 miles	-	2	12
16 to 20 miles	-	4	10
21 to 50 miles	-	10	13
51 to 100 miles	-	14	10
101 to 500 miles	-	27	10
501 to 1000 miles	-	11	7
1001 miles or more	-	26	13
Median(miles)	-	305	25

Table 2: Summary Statistics

We report the summary statistics of the Survey of Consumer Finance for 1995, 1998, 2001, 2004, 2007, and 2010 waves. % INVESTMENT REAL ESTATE IN TOTAL ASSETS is the fraction of investment real estate in household's total asset. % INVESTMENT REAL ESTATE IN REAL ESTATE ASSETS is the fraction of investment real estate in household's total real estate asset. % VACATION HOME IN TOTAL ASSETS is the fraction of vacation home in household's total asset. % DIRECTLY-HELD STOCKS IN TOTAL ASSETS is the fraction of stock holdings in household's total asset. HAVE INVESTMENT REAL ESTATE is the dummy variable for having investment real estate. HAVE VACATION HOME is the dummy variable for having vacation home. HAVE DIRECTLY-HELD STOCKS is the dummy variable for having directly-held stocks. Log RATIO is the log ratio of the total book market value of firms headquartered in a MSA to the income in that MSA. UNEMP is unemployment rate from the Bureau of Labor Statistics. HOME PRICE INDEX is the Federal Housing Finance Agency (FHFA) Housing Price Index at the MSA level. PAST PRICE APPRECIATION is the annual past home price appreciation in an MSA using the FHFA House Price Index, where $\frac{HPI_t - HPI_{t-1}}{HPI_{t-1}}$. Elasticity is the housing supply elasticity by MSA from Saiz (2010). Housing Affordability Index Backfilled is the Housing Affordability Index from the National Association of Realtors. Since the index is available only for 2009 to 2011, we backfilled the data using time-series average of available Housing Affordability Index. FAMILY SIZE is the number of people in each Primary Economic Unit. Log HOUSEHOLD INCOME is the log of household income. AGE, RACE, EDUCATION, and FAMILY STRUCTURE are reported. LWP stands for the "living with partner" in the FAMILY STRUCTURE variable. NET WORTH CATEGORY is the dummy variable on quintiles for household net worth amount. We report summary statistics for each group. UNATTACHED FEMALE is the gender variable for Primary Economic Unit head, which is equal to 1 for un-partnered females and zero for everybody else.

Variable		Obs	Mean	Std.	Max	Min
% INVESTMENT REAL ESTATE IN TOTAL ASSETS		110890	0.03	0.09	1.00	0
% INVESTMENT REAL ESTATE IN REAL ESTATE ASSETS		82878	0.08	0.20	1	0
% VACATION HOME IN TOTAL ASSETS		110890	0.01	0.04	0.87	0
% DIRECTLY-HELD STOCKS IN TOTAL ASSETS		108327	0.02	0.08	1	0
HAVE INVESTMENT REAL ESTATE		110890	0.13	0.34	1	0
HAVE VACATION HOME		110890	0.06	0.23	1	0
HAVE DIRECTLY-HELD STOCKS		110890	0.21	0.40	1	0
log RATIO		110890	-1.53	1.29	0.97	-7.34
UNEMP		110890	5.54	2.25	14.33	1.75
HOME PRICE INDEX		110890	160	52.09	323	100
PAST PRICE APPRECIATION		110890	0.04	0.06	0.29	-0.13
Elasticity		108235	1.59	0.84	5.36	0.63
Housing Affordability Index Backfilled		99820	122	55.68	383	51.17
FAMILY SIZE		110890	2.43	1.40	13.00	1
log HOUSEHOLD INCOME		110890	10.84	1.11	19.64	0
AGE	<35	110890	0.23	0.42	1	0
	35 – 49	110890	0.33	0.47	1	0
	50 – 64	110890	0.24	0.43	1	0
	≥65	110890	0.20	0.40	1	0
RACE	White	110890	0.73	0.44	1	0
	Black	110890	0.14	0.35	1	0
	Hispanic	110890	0.09	0.29	1	0
	Asian & Other	110890	0.04	0.20	1	0
EDUCATION	Less than High School	110890	0.13	0.34	1	0
	High School or GED only	110890	0.29	0.45	1	0
	Some college (or Associate's), but no Bachelor's	110890	0.25	0.43	1	0
	Bachelor's	110890	0.20	0.40	1	0
	MA or MS(non-MBA), plus nursing degree	110890	0.07	0.26	1	0
	Advanced (PhD, JD, MBA, MD)	110890	0.06	0.24	1	0
FAMILY STRUCTURE	Not Married/LWP+Children	110890	0.12	0.32	1	0
	Not Married/LWP+no Children+Head(<55)	110890	0.16	0.37	1	0
	Not Married/LWP+no Children+Head(≥55)	110890	0.13	0.34	1	0
	Married/LWP+Children	110890	0.32	0.47	1	0
	Married/LWP+no Children	110890	0.26	0.44	1	0
NET WORTH	CATEGORY 1/5 (Bottom)	110890	-4018	39137	15685	-2.74e+07
	CATEGORY 2/5	110890	49427	26695	126363	8298
	CATEGORY 3/5	110890	193224	72139	390303	77280
	CATEGORY 4/5	110890	521158	173824	959610	227826
	CATEGORY 5/5 (Top)	110890	3472319	1.05e+07	4.80e+09	539836
UNATTACHED FEMALE		110890	0.27	0.44	1	0

Table 3: Investment Real Estate, Vacation Homes, and Directly-held Stocks Ownership in the Survey of Consumer Finances

We report the household-level Logit regression results of investment home ownership, vacation home ownership, and directly-held stocks ownership on log RATIO. Column (1) reports the Logit regression result of investment real estate ownership (HAVE INVESTMENT REAL ESTATE) on log RATIO with household-level and MSA-level control variables. Household level control variables include EDUCATION, FAMILY STRUCTURE, RACE, UNATTACHED FEMALE, AGE, NET WORTH CATEGORY, FAMILY SIZE, and log HOUSEHOLD INCOME. MSA-level control variables include unemployment rate (UNEMP), HOME PRICE INDEX, Housing Affordability Index Backfilled, Elasticity, and PAST PRICE APPRECIATION. Column (2) reports the Logit regression result of vacation home ownership (HAVE VACATION HOME) on log RATIO with controls. Column (3) reports the Logit regression result of directly-held stock ownership (HAVE DIRECTLY-HELD STOCKS) on log RATIO with controls. EDUCATION, FAMILY STRUCTURE, and RACE are not reported for brevity. The entries in the brackets are the marginal probabilities that household does not own the asset. The table reports point estimates with t-statistics in parentheses. All the standard errors are clustered by MSA. ***, **, * denotes 1%, 5%, and 10% statistical significance.

Variables	(1) HAVE INVESTMENT REAL ESTATE	(2) HAVE VACATION HOME	(3) HAVE DIRECTLY-HELD STOCKS
log RATIO	-0.0920**	-0.0295	0.0776***
	(-2.520)	(-0.605)	(3.452)
	[-0.0087]	[-0.0015]	[0.0098]
Elasticity	0.0844	-0.0692	-0.0124
	(1.365)	(-0.774)	(-0.204)
	[0.008]	[-0.0035]	[-0.0016]
UNEMP	-0.00582	-0.124***	-0.0177
	(-0.185)	(-2.795)	(-0.579)
	[-0.0006]	[-0.0062]	[-0.0022]
HOME PRICE INDEX	-0.000703	-0.000937	-0.00153
	(-0.540)	(-0.558)	(-1.049)
	[-0.00007]	[-0.00005]	[-0.00019]
PAST PRICE APPRECIATION	-0.527	-0.647	-0.742
	(-0.750)	(-0.846)	(-0.959)
	[-0.05001]	[-0.0324]	[-0.0936]
Housing Affordability Index Backfilled	0.000560	0.000482	0.000596
	(0.596)	(0.754)	(1.487)
	[0.00005]	[0.00002]	[0.00008]
UNATTACHED FEMALE	-0.293**	0.321	-0.231***
	(-2.499)	(1.607)	(-2.938)
	[-0.0278]	[0.0161]	[-0.0291]
FAMILY SIZE	0.0638	-0.209	-0.121
	(0.389)	(-1.322)	(-0.893)
	[0.00604]	[-0.0105]	[-0.0153]
FAMILY SIZE^2	-0.00969	0.0202	0.00338
	(-0.544)	(1.070)	(0.197)
	[-0.0009]	[0.001]	[0.0004]
log HOUSEHOLD INCOME	0.252***	0.432***	0.382***
	(6.334)	(9.247)	(8.180)
	[0.0239]	[0.0216]	[0.0482]
NET WORTH CATEGORY (2/5)	1.732***	1.255***	0.734***
	(7.804)	(4.014)	(5.838)
	[0.1643]	[0.0629]	[0.0927]
NET WORTH CATEGORY (3/5)	2.531***	1.878***	1.376***
	(11.19)	(6.808)	(13.01)
	[0.2401]	[0.094]	[0.1736]
NET WORTH CATEGORY (4/5)	3.370***	2.360***	2.066***
	(14.53)	(7.658)	(17.21)
	[0.3198]	[0.1182]	[0.2606]
NET WORTH CATEGORY (5/5)	4.224***	2.773***	2.669***
	(17.56)	(8.278)	(16.27)
	[0.4007]	[0.1389]	[0.3368]
Constant	-7.830***	-10.11***	-6.983***
	(-14.52)	(-14.01)	(-11.32)
Observations	99,255	99,255	99,255
R^2	0.206	0.166	0.229
Additional Demographic Controls	EDUCATION, RACE, FAMILY STRUCTURE, and AGE		
Year FE	Yes	Yes	Yes

Table 4: % Share of Investment Real Estate, Vacation Homes, and Directly-held Stocks in the Survey of Consumer Finances

We report the household-level Tobit regression results of % share of investment real estate, vacation home, and stock holdings in households' asset on log RATIO. Column (1) reports the Tobit regression result of % share of investment real estate in total assets (% INVESTMENT REAL ESTATE IN TOTAL ASSETS) on log RATIO with household-level and MSA-level control variables. Household level control variables include EDUCATION, FAMILY STRUCTURE, RACE, AGE, UNATTACHED FEMALE, FAMILY SIZE, NET WORTH CATEGORY, and log HOUSEHOLD INCOME. MSA-level control variables include unemployment rate (UNEMP), HOME PRICE INDEX, Housing Affordability Index Backfilled, Elasticity, and PAST PRICE APPRECIATION. Column (2) reports the Tobit regression result of % share of investment real estate in total real estate assets (% INVESTMENT REAL ESTATE IN REAL ESTATE ASSETS) on log RATIO with controls. Column (3) reports the Tobit regression result of % share of vacation home in total assets (% VACATION HOME IN TOTAL ASSETS) on log RATIO with controls. Column (4) reports the Tobit regression result of % directly-held stocks in total assets (%DIRECTLY-HELD STOCKS IN TOTAL ASSETS) on log RATIO with controls. EDUCATION, FAMILY STRUCTURE, AGE, RACE, and year FE are not reported for brevity. The table reports point estimates with t-statistics in parentheses. All the standard errors are clustered by MSA. ***, **, * denotes 1%, 5%, and 10% statistical significance.

Variables	(1) % INVESTMENT REAL ESTATE IN TOTAL ASSETS	(2) % INVESTMENT REAL ESTATE IN REAL ESTATE ASSETS	(3) % VACATION HOME IN TOTAL ASSETS	(4) % DIRECTLY-HELD STOCKS IN TOTAL ASSETS
log RATIO	-0.0107*	-0.0167	-0.00151	0.00781**
	(-1.652)	(-1.249)	(-0.299)	(2.343)
Econ Sig	-0.15	-0.11	-0.05	0.13
Elasticity	0.0102	0.0198	-0.00419	0.00150
	(0.962)	(0.850)	(-0.495)	(0.265)
UNEMP	0.000377	0.00196	-0.0119***	-0.000717
	(0.0761)	(0.179)	(-3.158)	(-0.246)
HOME PRICE INDEX	0.000177	0.000233	-4.82e-05	-0.000101
	(0.787)	(0.482)	(-0.298)	(-0.807)
PAST PRICE APPRECIATION	-0.0501	-0.0821	-0.0369	-0.00572
	(-0.579)	(-0.451)	(-0.521)	(-0.0986)
Housing Affordability Index Backfilled	6.04e-05	-2.16e-05	4.89e-05	6.36e-05
	(0.276)	(-0.0492)	(0.573)	(0.751)
UNATTACHED FEMALE	-0.0301*	-0.0891***	0.0168	-0.0194***
	(-1.816)	(-2.963)	(1.263)	(-2.836)
FAMILY SIZE	-0.00505	-0.0152	-0.0158	-0.0181*
	(-0.252)	(-0.444)	(-1.185)	(-1.856)
FAMILY SIZE^2	0.000483	0.00118	0.00120	0.00102
	(0.213)	(0.306)	(0.749)	(0.865)
log HOUSEHOLD INCOME	-0.00249	-0.000351	-0.00846*	0.0140***
	(-0.493)	(-0.0318)	(-1.693)	(3.421)
NET WORTH CATEGORY (2/5)	0.190***	0.0603	0.0929***	0.0533***
	(5.806)	(1.214)	(4.359)	(4.763)
NET WORTH CATEGORY (3/5)	0.276***	0.122***	0.129***	0.0944***
	(7.571)	(2.600)	(6.074)	(8.064)
NET WORTH CATEGORY (4/5)	0.355***	0.258***	0.152***	0.134***
	(8.069)	(4.598)	(6.265)	(8.697)
NET WORTH CATEGORY (5/5)	0.406***	0.391***	0.135***	0.179***
	(7.166)	(5.965)	(4.350)	(8.598)
Constant	-0.488***	-0.425*	-0.236**	-0.316***
	(-4.176)	(-1.950)	(-2.457)	(-4.606)
Observations	99,255	74,263	99,255	96,900
Additional Demographic Controls	EDUCATION, RACE, FAMILY STRUCTURE, and AGE			
Year FE	Yes	Yes	Yes	Yes

Table 5: SCF Results by Zip-code Income and by MSA Subprime Share

Panel A reports the regression results with the indicator variables for high income zip-code, the dummy variable for post year 2002, and their interactions. I_High Income Zipcode is an indicator variable for the top 1/3 zip-code in terms of average zip-code income, measured by the American Community Survey from 2006 to 2010. I_after2002 is an indicator variable for the years after 2002. Column (1) reports the household-level regression of investment home ownership (HAVE INVESTMENT REAL ESTATE) on log RATIO with interactions with the indicator variable for high income zip-code and the dummy variable for post year 2002, with household-level and MSA-level control variables. Household level control variables include EDUCATION, FAMILY STRUCTURE, RACE, UNATTACHED FEMALE, AGE, NET WORTH CATEGORY, and FAMILY SIZE. MSA-level control variables include unemployment rate (UNEMP) and PAST PRICE APPRECIATION. Column (2) reports the household-level Tobit regression of % share of investment real estate in total assets (% INVESTMENT REAL ESTATE IN TOTAL ASSETS) on log RATIO with interactions with the indicator variable for high income zip-code and the dummy variable for post year 2002, with household-level and MSA-level control variables. Panel B reports the regression results with the indicator variables for high subprime MSAs, the dummy variable for post year 2002, and their interactions. I_Subprime MSA is an indicator variable for the top 1/3 MSAs in terms of average subprime mortgage share among total mortgage origination from 2002 to 2007. Column (1) reports the household-level regression of HAVE INVESTMENT REAL ESTATE on log RATIO with interactions with the indicator variable for high subprime MSA and the dummy variable for post year 2002, with household-level and MSA-level control variables. Column (2) reports the household-level Tobit regression of % INVESTMENT REAL ESTATE IN TOTAL ASSETS on log RATIO with interactions with the indicator variable for high subprime MSA and the dummy variable for post year 2002, with household-level and MSA-level control variables. Some controls are not reported for brevity. The table reports point estimates with t-statistics in parentheses. All the standard errors in columns are clustered by MSA. ***, **, * denotes 1%, 5%, and 10% statistical significance.

Panel A: *By Zip-Code Income*

Variables	(1) HAVE INVESTMENT REAL ESTATE	(2) % INVESTMENT REAL ESTATE IN TOTAL ASSETS
log RATIO * I_High Income Zipcode * I_after2002	0.0107	0.00152
	(1.114)	(0.578)
log RATIO	-0.00373	-0.000490
	(-0.691)	(-0.427)
log RATIO * I_High Income Zipcode	-0.00623	-0.00181
	(-0.996)	(-0.974)
log RATIO * I_after2002	-0.00964	-0.00209
	(-1.337)	(-1.039)
I_High Income Zipcode * I_after2002	0.0356**	0.00744
	(2.043)	(1.527)
I_High Income Zipcode	-0.0352***	-0.00963***
	(-2.805)	(-2.814)
I_after2002	-0.00866	-0.00243
	(-0.652)	(-0.589)
Constant	0.0233	0.000897
	(0.857)	(0.107)
Observations	111,375	111,375
R^2	0.164	0.066

Panel B: *By MSA Subprime Share*

Variables	(1) HAVE INVESTMENT REAL ESTATE	(2) % INVESTMENT REAL ESTATE IN TOTAL ASSETS
log RATIO * I_Subprime MSA * I_after2002	-0.0228***	-0.00445**
	(-3.188)	(-2.181)
log RATIO	-0.0127**	-0.00259**
	(-2.424)	(-2.447)
log RATIO * I_Subprime MSA	0.0125	0.00238
	(1.395)	(1.530)
log RATIO * I_after2002	0.00576	0.000615
	(0.980)	(0.446)
I_Subprime MSA * I_after2002	-0.0373***	-0.00660
	(-2.637)	(-1.628)
I_Subprime MSA	0.0225	0.00467
	(1.043)	(1.332)
I_after2002	0.0311***	0.00574*
	(2.812)	(1.898)
Constant	0.00394	-0.00342
	(0.134)	(-0.403)
Observations	111,375	111,375
R^2	0.164	0.066

Table 6: SCF Results by County Subprime Share

We report the regression results with the indicator variables for high subprime counties, the dummy variable for post year 2002, and their interactions. I_Subprime County is an indicator variable for the top 1/3 counties (within MSAs) in terms of average subprime mortgage share among total mortgage origination from 2002 to 2007. Column (1) reports the household-level regression of HAVE INVESTMENT REAL ESTATE on log RATIO with interactions with the indicator variable for high subprime county and the dummy variable for post year 2002, with household-level and MSA-level control variables. Household level control variables include EDUCATION, FAMILY STRUCTURE, RACE, UNATTACHED FEMALE, AGE, NET WORTH CATEGORY, and FAMILY SIZE. MSA-level control variables include unemployment rate (UNEMP) and PAST PRICE APPRECIATION. Column (2) reports the household-level Tobit regression of % INVESTMENT REAL ESTATE IN TOTAL ASSETS on log RATIO with interactions with the indicator variable for high subprime county and the dummy variable for post year 2002, with household-level and MSA-level control variables. Some controls are not reported for brevity. The table reports point estimates with t-statistics in parentheses. All the standard errors in columns are clustered by MSA. ***, **, * denotes 1%, 5%, and 10% statistical significance.

Variables	(1) HAVE INVESTMENT REAL ESTATE	(2) % INVESTMENT REAL ESTATE IN TOTAL ASSETS
log RATIO * I_Subprime County * I_after2002	-0.0256**	-0.00636**
	(-2.516)	(-2.195)
log RATIO	-0.00460	-0.00112
	(-0.836)	(-1.065)
log RATIO * I_Subprime County	0.0239**	0.00503***
	(2.442)	(2.905)
log RATIO * I_after2002	-0.00175	-0.000909
	(-0.287)	(-0.670)
I_Subprime County * I_after2002	-0.0470**	-0.00455
	(-2.314)	(-0.563)
I_Subprime County	0.0465*	0.00739*
	(1.955)	(1.710)
I_after2002	-0.00344	0.00557**
	(-0.337)	(2.279)
Constant	-0.0753***	-0.00300
	(-2.635)	(-0.363)
Observations	111,456	111,456
R^2	0.061	0.020

Table 7: Summary Statistics

We report the summary statistics of MSA-level variables. NON-OWNER OCCUPIED HOME PURCHASE AMTO is the amount of non-owner occupied mortgage originations for home purchase in an MSA from the Home Mortgage Disclosure Act (HMDA). TOTAL HOME PURCHASE AMTO is the total amount of mortgage originations for home purchase in an MSA from the HMDA. %NON-OCCUPIED MORTGAGES is the ratio of NON-OWNER OCCUPIED HOME PURCHASE AMTO to TOTAL HOME PURCHASE AMTO. Log RATIO is the log ratio of the total book value of firms headquartered in a MSA to the income in that MSA as in Hong, Kubik, and Stein (2008). Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log. BOOK PER CAPITA is the per capita book value of firms headquartered in a MSA from the COMPUSTAT. INCOME PER CAPITA is the per capita income from the Bureau of Economic Analysis. PAST PRICE APPRECIATION is the annual home price appreciation from Federal Housing Finance Agency (FHFA). logPOP is the log number of population from the BEA. UNEMP is unemployment rate from the Bureau of Labor Statistics. Housing supply elasticity (Elasticity) is from Saiz (2010). Housing Affordability Index is from the National Association of Realtors and is available from 2009 to 2011. Log PR is the log of price to rent ratio using the FHFA home price index and the fair market rent from the U.S. Department of Housing and Urban Development.

Variables	Obs	Mean	Std	Max	Min
NON-OWNER OCCUPIED HOME PURCHASE AMTO (millions)	5082	198	603	11845	1
TOTAL HOME PURCHASE AMTO (millions)	5082	1670	5244	84397	7
%NON-OCCUPIED MORTGAGES	5082	0.14	0.1	0.89	0.01
log RATIO	5082	-6.33	4.68	2.4	-11.51
BOOK PER CAPITA	5082	5355	16474	254821	0
INCOME PER CAPITA	5082	31463	7222	80139	12723
PAST PRICE APPRECIATION (%)	5079	3.18	6.65	34.8	-37.9
logPOP	5082	12.63	1.06	16.76	10.75
UNEMP	5082	5.92	2.78	30.06	1.23
Elasticity	3920	2.52	1.42	12.15	0.63
Housing Affordability Index	444	130.60	38.83	383.4	35.1
log PR	8684	-1.5	0.23	-0.71	-2.4

Table 8: Baseline Regressions

We report the baseline regression results of the non-owner-occupied mortgage originations on RATIO. Left-hand side variable (log %NON-OCCUPIED MORTGAGES) is the log ratio of non-owner-occupied mortgage originations for home purchase to total mortgage origination for home purchase. Log RATIO is the log ratio of the total book market value of firms headquartered in a MSA to the income in that MSA. Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log. Columns (1)-(3) report the panel regression using MSA-level samples from 1998 to 2011. Column (1) reports the univariate regression. Columns (2)-(3) report the multivariate regression results. Controls include unemployment rate (UNEMP) and housing supply elasticity (Elasticity) from Saiz (2010). Columns (4)-(6) report the cross-sectional regression by averaging observations across the sample period. Column (4) reports the univariate regression. Columns (5)-(6) report the multivariate regression results. The table reports point estimates with t-statistics in parentheses. All the standard errors in columns (1)-(3) are clustered by MSA. ***, **, * denotes 1%, 5%, and 10% statistical significance.

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	log %NON-OCCUPIED MORTGAGES					
log RATIO	-0.0366***	-0.0328***	-0.0401***	-0.0603***	-0.0523***	-0.0526***
	(-7.481)	(-6.624)	(-5.869)	(-4.406)	(-3.680)	(-2.849)
Econ Sig	-0.28	-0.25	-0.29	-0.23	-0.20	-0.18
UNEMP		0.0112	0.00299		0.0243**	0.0277*
		(1.382)	(0.338)		(1.984)	(1.844)
Elasticity			-0.0813***			-0.0591***
			(-3.861)			(-2.839)
Constant	-2.365***	-2.900***	-2.700***	-2.691***	-2.766***	-2.673***
	(-66.02)	(-59.67)	(-34.21)	(-21.68)	(-21.40)	(-17.44)
Observations	5,082	5,082	3,920	363	363	280
Adj R^2	0.078	0.304	0.325	0.048	0.056	0.072
Year FE	No	Yes	Yes	No	No	No
Sample Period	1998-2011	1998-2011	1998-2011	Average CX	Average CX	Average CX
Area Unit	MSA	MSA	MSA	MSA	MSA	MSA

Table 9: Decomposing the RATIO

We report the regression results decomposing the RATIO variable. Left-hand side variable (log %NON-OCCUPIED MORTGAGES) is the log ratio of non-owner-occupied mortgage originations for home purchase to total mortgage origination for home purchase. Right-hand side variables are the log of book value per capita (log BOOK PER CAPITA) and the log of per capital income (log INCOME PER CAPITA). For the MSAs with BOOK PER CAPITA equal to zero, we assign one dollar of total book value to the MSAs before we take log. Columns (1)-(2) report the panel regression using MSA-level samples from 1998 to 2011. Control variables include unemployment rate (UNEMP) and housing supply elasticity from Saiz (2010). Columns (3)-(4) report the cross-sectional regression by averaging observations across the sample period. The table reports point estimates with t-statistics in parentheses. All the standard errors in columns (1)-(2) are clustered by MSA. ***, **, * denotes 1%, 5%, and 10% statistical significance.

	(1)	(2)	(3)	(4)
Variables	log %NON-OCCUPIED MORTGAGES			
log BOOK PER CAPITA	-0.0129***	-0.0147***	-0.0159***	-0.0186***
	(-4.549)	(-4.046)	(-4.812)	(-4.595)
Econ Sig	-0.21	-0.22	-0.28	-0.30
log INCOME PER CAPITA	-0.217	-0.536**	-0.212	-0.567***
	(-1.015)	(-2.189)	(-1.185)	(-2.678)
Econ Sig	-0.08	-0.19	-0.07	-0.19
UNEMP	0.00774	-0.00967	0.0128	-0.00446
	(0.795)	(-0.852)	(1.047)	(-0.289)
Elasticity		-0.0989***		-0.101***
		(-4.487)		(-4.688)
Constant	-0.500	3.051	-0.0623	3.909*
	(-0.230)	(1.217)	(-0.0338)	(1.770)
Observations	5,082	3,920	363	280
Adj R^2	0.298	0.326	0.110	0.163
Year FE	Yes	Yes	No	No
Sample Period	1998-2011	1998-2011	Average CX	Average CX
Area Unit	MSA	MSA	MSA	MSA

Table 10: Robustness Check

We report robustness checks for the baseline regression. Left-hand side variable (log %NON-OCCUPIED MORTGAGES) is the log ratio of non-owner-occupied mortgage originations for home purchase to total mortgage origination for home purchase. Log RATIO is the log ratio of the total book market value of firms headquartered in a MSA to the income in that MSA. Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log. Column (1) reports the regression result with log number of population (logPOP) and other control variables. Control variables include unemployment rate (UNEMP) and housing supply elasticity (Elasticity) from Saiz (2010). Column (2) reports the regression result without vacation MSAs. Vacation MSAs are the popular vacation home areas listed from CNBC, Barron, and Forbes. Column (3) reports the regression result with the future annual home price appreciation (FUTURE PRICE APPRECIATION), where $\frac{HPI_{t+1}-HPI_t}{HPI_t}$. Column (4) reports the regression result with the past annual home price appreciation (PAST PRICE APPRECIATION), where $\frac{HPI_t-HPI_{t-1}}{HPI_{t-1}}$. In column (5)-(8), we control the Housing Affordability Index from the National Association of Realtors. In column (5)-(6), we control the Housing Affordability Index which is available from 2009 to 2011. In columns (7)-(8), we control for the Housing Affordability Index Backfilled, which is the time-series average of the Housing Affordability Index for 1998 to 2011 period. The table reports point estimates with t-statistics in parentheses. All the standard errors are clustered by MSA. ***, **, * denotes 1%, 5%, and 10% statistical significance.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	log %NON-OCCUPIED MORTGAGES							
log RATIO	-0.0275***	-0.0391***	-0.0396***	-0.524***	-0.0304***	-0.0443***	-0.0301***	-0.0466***
	(-3.833)	(-6.048)	(-5.873)	(-4.077)	(-3.184)	(-3.458)	(-3.582)	(-4.152)
Econ Sig	-0.210	-0.30	-0.29	-0.22	-0.22	-0.28	-0.19	-0.26
logPOP	-0.107***							
	(-3.609)							
FUTURE PRICE APPRECIATION			1.256***					
			(6.165)					
PAST PRICE APPRECIATION				1.072***				
				(4.386)				
Housing Affordability Index					-0.00309***	-0.00192		
					(-2.941)	(-1.360)		
Housing Affordability Index Backfilled							-0.00226***	-0.00106
							(-2.641)	(-0.914)
Elasticity	-0.107***	-0.0584***	-0.0763***	-0.0533***		-0.0714		-0.101***
	(-4.261)	(-3.004)	(-3.763)	(-2.794)		(-1.639)		(-2.700)
UNEMP	0.00707	-0.00228	0.00762	0.0236**	0.0791***	0.0592**	0.0206	-0.000886
	(0.772)	(-0.272)	(0.849)	(2.460)	(3.788)	(2.579)	(1.146)	(-0.0431)
Constant	-1.226***	-2.773***	-2.778***	-2.660***	-2.462***	-2.302***	-2.700***	-2.586***
	(-2.844)	(-35.78)	(-34.92)	(-29.44)	(-11.73)	(-10.28)	(-21.78)	(-18.33)
Observations	3,920	3,584	3,920	3,920	444	417	2,156	2,030
Adj R^2	0.346	0.350	0.334	0.302	0.156	0.169	0.316	0.349
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Period	1998-2011	1998-2011	1998-2011	1998-2011	2009-2011	2009-2011	1998-2011	1998-2011
Area Unit	MSA	MSA	MSA	MSA	MSA	MSA	MSA	MSA

Table 11: Owner-Occupied Originations

We report the regression results of the owner-occupied mortgage originations on RATIO. Left-hand side variable (log OWNER OCCUPIED MORTGAGES) is the log amount of owner-occupied mortgage originations for home purchase. Log RATIO is the log ratio of the total book market value of firms headquartered in a MSA to the income in that MSA. Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log. Columns (1)-(3) report the panel regression using MSA-level samples from 1998 to 2011. Column (1) reports the univariate regression. Columns (2)-(3) report the multivariate regression results. Control variables include the log amount of total mortgage originations for home purchase (logAMTO), unemployment rate (UNEMP), and housing supply elasticity (Elasticity) from Saiz (2010). Columns (4)-(6) report the cross-sectional regression by averaging observations across the sample period. Column (4) reports the univariate regression. Columns (5)-(6) report the multivariate regression results. The table reports point estimates with t-statistics in parentheses. All the standard errors in columns (1)-(3) are clustered by MSA. ***, **, * denotes 1%, 5%, and 10% statistical significance.

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	log OWNER OCCUPIED MORTGAGES					
log RATIO	0.172***	0.0100***	0.0106***	0.299***	0.0140***	0.0151***
	(14.39)	(4.311)	(3.324)	(8.365)	(3.292)	(2.833)
Econ Sig	0.55	0.03	0.03	0.40	0.02	0.02
logAMTO		0.980***	0.992***		0.991***	1.005***
		(139.4)	(151.0)		(178.1)	(137.4)
UNEMP		-0.00497*	-0.00279		-0.00700**	-0.00641
		(-1.775)	(-0.953)		(-2.055)	(-1.497)
Elasticity			0.0214***			0.0228***
			(3.697)			(3.302)
Constant	13.77***	0.245**	0.0203	15.44***	0.114	-0.109
	(123.6)	(2.384)	(0.210)	(47.59)	(1.223)	(-0.865)
Observations	5,082	5,082	3,920	363	363	280
Adj R^2	0.302	0.991	0.991	0.160	0.991	0.990
Year FE	No	Yes	Yes	No	No	No
Sample Period	1998-2011	1998-2011	1998-2011	Average CX	Average CX	Average CX
Area Unit	MSA	MSA	MSA	MSA	MSA	MSA

Table 12: HMDA Results by MSA Income and by MSA Subprime Share

We report the regression results of non-owner-occupied mortgage originations on RATIO with interactions with the indicator variable for high income MSAs/high subprime share MSAs and the dummy variable for post year 2002. Left-hand side variable (log %NON-OCCUPIED MORTGAGES) is the log ratio of non-owner-occupied mortgage originations for home purchase to total mortgage origination for home purchase. Log RATIO is the log ratio of the total book market value of firms headquartered in a MSA to the income in that MSA. Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log. In panel A, we report the regression results with the indicator variables for high income MSA, the dummy variable for post year 2002, and their interactions. I_High Income MSA is an indicator variable for the top 1/3 MSAs in terms of average income per capita from 1998 to 2011. I_after2002 is an indicator variable for the years after 2002. In panel B, we report the regression results with the indicator variables for high subprime share MSAs, the dummy variable for post year 2002, and their interactions. I_Subprime MSA is an indicator variable for the top 1/3 MSAs in terms of average subprime mortgage share among total mortgage origination from 2002 to 2007. I_after2002 is an indicator variable for the years after 2002. We also control variables include unemployment rate (UNEMP) and housing supply elasticity (Elasticity) from Saiz (2010). The table reports point estimates with t-statistics in parentheses. All the standard errors in columns are clustered by MSA. ***, **, * denotes 1%, 5%, and 10% statistical significance.

Panel A: *By MSA per capita Income*

Variables	(1) log %NON-OCCUPIED MORTGAGES
log RATIO * I_High Income MSA * I_after2002	0.0325**
	(2.012)
log RATIO	-0.0196**
	(-2.450)
log RATIO * I_High Income MSA	-0.0975***
	(-3.265)
log RATIO * I_after2002	-0.00216
	(-0.487)
I_High Income MSA * I_after2002	0.0381
	(0.814)
I_High Income MSA	-0.292***
	(-3.607)
I_after2002	0.626***
	(18.58)
Elasticity	-0.0774***
	(-3.732)
UNEMP	0.0118
	(1.318)
Constant	-2.633***
	(-29.53)
Observations	3,920
Adjusted R^2	0.366
Year FE	Yes
Sample Period	1998-2011
Area Unit	MSA

Panel B: *By MSA Subprime Share*

Variables	(2) log %NON-OCCUPIED MORTGAGES
log RATIO * I_Subprime MSA * I_after2002	-0.0131*
	(-1.692)
log RATIO	-0.0475***
	(-4.481)
log RATIO * I_Subprime MSA	0.0183
	(1.198)
log RATIO * I_after2002	0.00629
	(1.133)
I_Subprime MSA * I_after2002	-0.0224
	(-0.482)
I_Subprime MSA	0.0900
	(1.014)
I_after2002	0.637***
	(22.04)
Elasticity	-0.0792***
	(-3.807)
UNEMP	0.00243
	(0.223)
Constant	-2.733***
	(-30.51)
Observations	3,920
Adjusted R^2	0.326
Year FE	Yes
Sample Period	1998-2011
Area Unit	MSA

Table 13: Price to Rent Ratio

We regress price to rent ratio on RATIO. Left-hand side variable (log PR) is the log of price to rent ratio. We use the home price index from the Federal Housing Finance Agency and the fair market rent from the U.S. Department of Housing and Urban Development. Log RATIO is the log ratio of the total book market value of firms headquartered in a MSA to the income in that MSA. Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log. Columns (1)-(3) report the panel regression using MSA-level samples from 1990 to 2011. Column (1) reports the univariate regression. Columns (2)-(3) report the multivariate regression results. Control variables include unemployment rate (UNEMP) and housing supply elasticity (Elasticity) from Saiz (2010). Columns (4)-(6) report the cross-sectional regression by averaging observations across the sample period. Column (4) reports the univariate regression. Columns (5)-(6) report the multivariate regression results. The table reports point estimates with t-statistics in parentheses. All the standard errors in columns (1)-(3) are clustered by MSA. ***, **, * denotes 1%, 5%, and 10% statistical significance.

	(1)	(2)	(3)	(4)	(5)	(6)
Variables			log PR			
log RATIO	-0.00400***	-0.0107***	-0.00780***	-0.0167***	-0.0179***	-0.0120***
	(-2.709)	(-5.876)	(-4.384)	(-7.515)	(-7.697)	(-4.531)
Econ Sig	-0.08	-0.22	-0.15	-0.37	-0.39	-0.25
UNEMP		-0.0105***	-0.00645***		-0.00656*	0.000365
		(-4.678)	(-2.684)		(-1.688)	(0.0869)
Elasticity			0.0478***			0.0493***
			(6.844)			(8.498)
Constant	-1.526***	-1.654***	-1.773***	-1.602***	-1.570***	-1.709***
	(-121.1)	(-99.22)	(-65.97)	(-97.49)	(-62.73)	(-59.00)
Observations	8,748	7,859	6,126	361	361	280
Adj R^2	0.006	0.339	0.425	0.134	0.138	0.308
Year FE	No	Yes	Yes	No	No	No
Sample Period	1986-2011	1990-2011	1990-2011	Average CX	Average CX	Average CX
Area Unit	MSA	MSA	MSA	MSA	MSA	MSA

Table 14: Annual Home Price Appreciation

We regress annual home price appreciation on RATIO. Left-hand side variable (HPIAPP) is the annual home price appreciation, defined as $\frac{HPI_t - HPI_{t-1}}{HPI_{t-1}}$ using the home price index (HPI) from the Federal Housing Finance Agency. Log RATIO is the log ratio of the total book market value of firms headquartered in a MSA to the income in that MSA. Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log. I_Subprime MSA is an indicator variable for the top 1/3 MSAs in terms of average subprime mortgage share among total mortgage origination from 2002 to 2007. Column (1) reports the panel regression using MSA-level samples from 1998 to 2005. Column (2) reports the panel regression from 2002 to 2005. Column (3) reports the panel regression from 2002 to 2007. Control variables include unemployment rate (UNEMP) and housing supply elasticity (Elasticity) from Saiz (2010). The table reports point estimates with t-statistics in parentheses. All the standard errors are clustered by MSA. ***, **, * denotes 1%, 5%, and 10% statistical significance.

	(1)	(2)	(3)
Variables	Annual Home Price Appreciation		
log RATIO * I_Subprime MSA	-0.00175**	-0.00294**	-0.00191*
	(-2.063)	(-2.184)	(-1.920)
log RATIO	-0.000134	-0.000771	-0.000882*
	(-0.308)	(-1.208)	(-1.925)
I_Subprime MSA	0.00566	0.00969	0.00984
	(1.052)	(1.131)	(1.594)
UNEMP	-0.00258**	-0.00389	-0.00598**
	(-2.556)	(-1.606)	(-2.569)
Elasticity	-0.0129***	-0.0192***	-0.0138***
	(-6.622)	(-6.747)	(-6.538)
Constant	0.0853***	0.161***	0.115***
	(12.18)	(10.56)	(7.872)
Observations	2,240	1,120	1,680
Adj R^2	0.296	0.320	0.306
Year FE	Yes	Yes	Yes
Sample Period	1998-2005	2002-2005	2002-2007
Area Unit	MSA	MSA	MSA

Table 15: Homeownership Rate

We regress homeownership rate on RATIO. Left-hand side variable is the homeownership rate, computed by dividing the number of owner-occupied housing units by the number of occupied housing units of households, using the Integrated Public Use Microdata Series. Log RATIO is the log ratio of the total book market value of firms headquartered in a MSA to the income in that MSA. Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log. Panel A reports the pooled panel regression using MSA-level samples from 2005 to 2011. Control variables include unemployment rate (UNEMP), housing supply elasticity (Elasticity) from Saiz (2010), and the log number of population (logPOP). Panel B reports the cross-sectional regression by averaging observations across the sample period. Panel C reports the Fama-MacBeth regressions from 2005 to 2011. The table reports point estimates with t-statistics in parentheses. The standard errors in Panel A are clustered by MSA. ***, **, * denotes 1%, 5%, and 10% statistical significance.

	Panel A: Pooled Panel		**Panel B:** Cross-Sectional Average		**Panel C:** Fama-MacBeth	
	(1)	(2)	(3)	(4)	(5)	(6)
Variables	log HOMEOWNERSHIP RATE		log HOMEOWNERSHIP RATE		log HOMEOWNERSHIP RATE	
log RATIO	0.00216	0.00420***	0.00215	0.00435**	0.00212***	0.00416***
	(1.558)	(2.650)	(1.451)	(2.541)	(4.558)	(8.156)
Econ Sig	0.11	0.20	0.11	0.22		
UNEMP	-0.00524*	-0.00476	-0.00606*	-0.00539*	-0.00576***	-0.00523***
	(-1.668)	(-1.546)	(-1.859)	(-1.666)	(-5.296)	(-4.773)
Elasticity	0.00737*	0.00266	0.00721	0.00251	0.00725***	0.00257***
	(1.708)	(0.552)	(1.573)	(0.512)	(10.92)	(4.562)
logPOP		-0.0176**		-0.0180**		-0.0176***
		(-2.228)		(-2.499)		(-14.91)
Constant	4.259***	4.507***	4.260***	4.513***	4.256***	4.503***
	(222.0)	(39.33)	(164.4)	(43.32)	(532.2)	(459.9)
Observations	1,638	1,638	234	234	1,638	1,638
R^2	0.073	0.095	0.044	0.070	0.045	0.068
Year FE	Yes	Yes	No	No	No	No
Sample Period	2005-2011	2005-2011	Average CX	Average CX	2005-2011	2005-2011
Area Unit	MSA	MSA	MSA	MSA	MSA	MSA

Figure 1: log %NON-OCCUPIED MORTGAGES and log RATIO

The figure plots log %NON-OCCUPIED MORTGAGES and logRATIO across MSAs. Log %NON-OCCUPIED MORTGAGES and log RATIO are constructed by averaging observations across the sample period. %NON-OCCUPIED MORTGAGES is the ratio of non-owner-occupied mortgage originations for home purchase to total mortgage origination for home purchase. Log RATIO is the log of ratio of the total book market value of firms headquartered in a MSA to the income in that MSA as in Hong, Kubik, and Stein (2008). Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log.

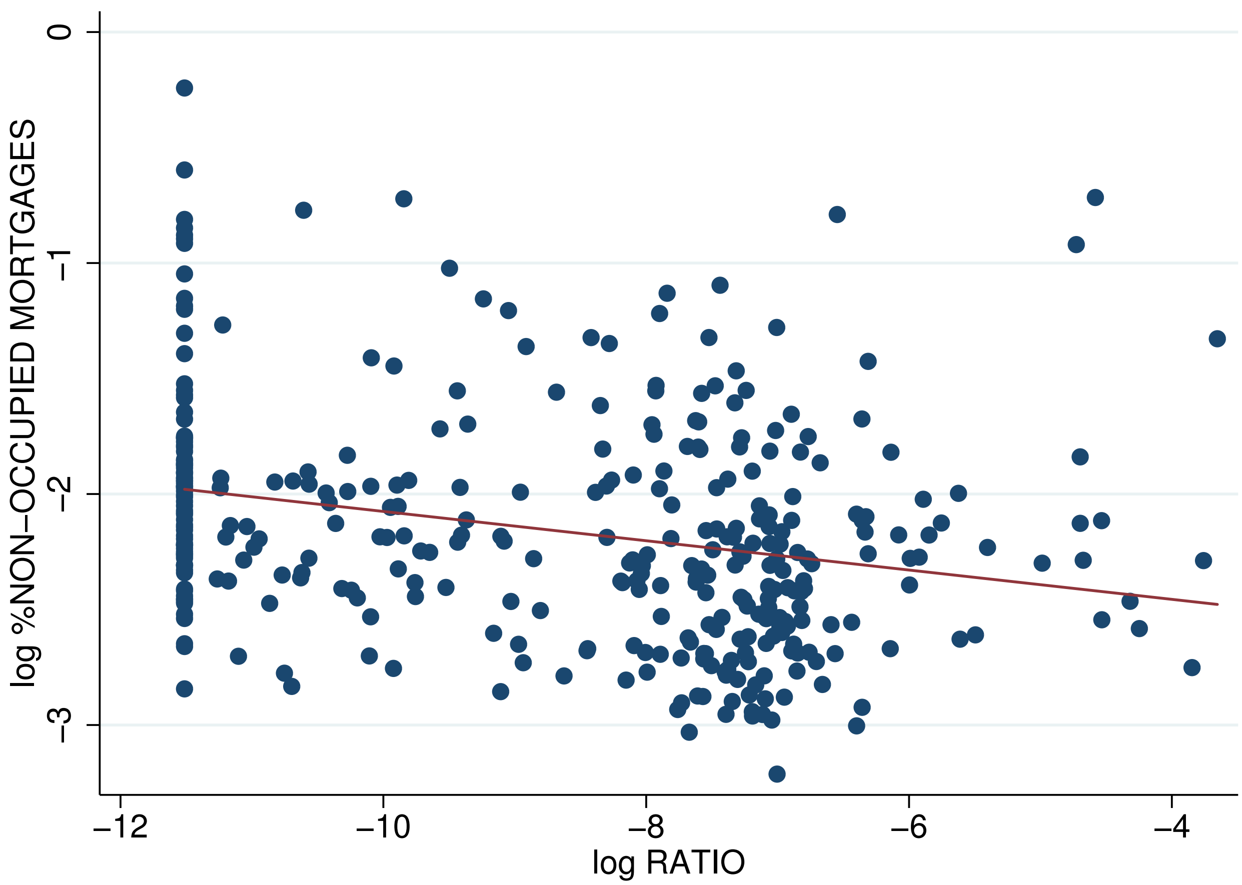

Figure 2: log PR and log RATIO

The figure plots log PR and logRATIO across MSAs. Log PR and log RATIO are constructed by averaging observations across the sample period. PR is the price to rent ratio. We use the home price index from the Federal Housing Finance Agency and the fair market rent from the U.S. Department of Housing and Urban Development. Log RATIO is the log of ratio of the total book market value of firms headquartered in a MSA to the income in that MSA as in Hong, Kubik, and Stein (2008). Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log.

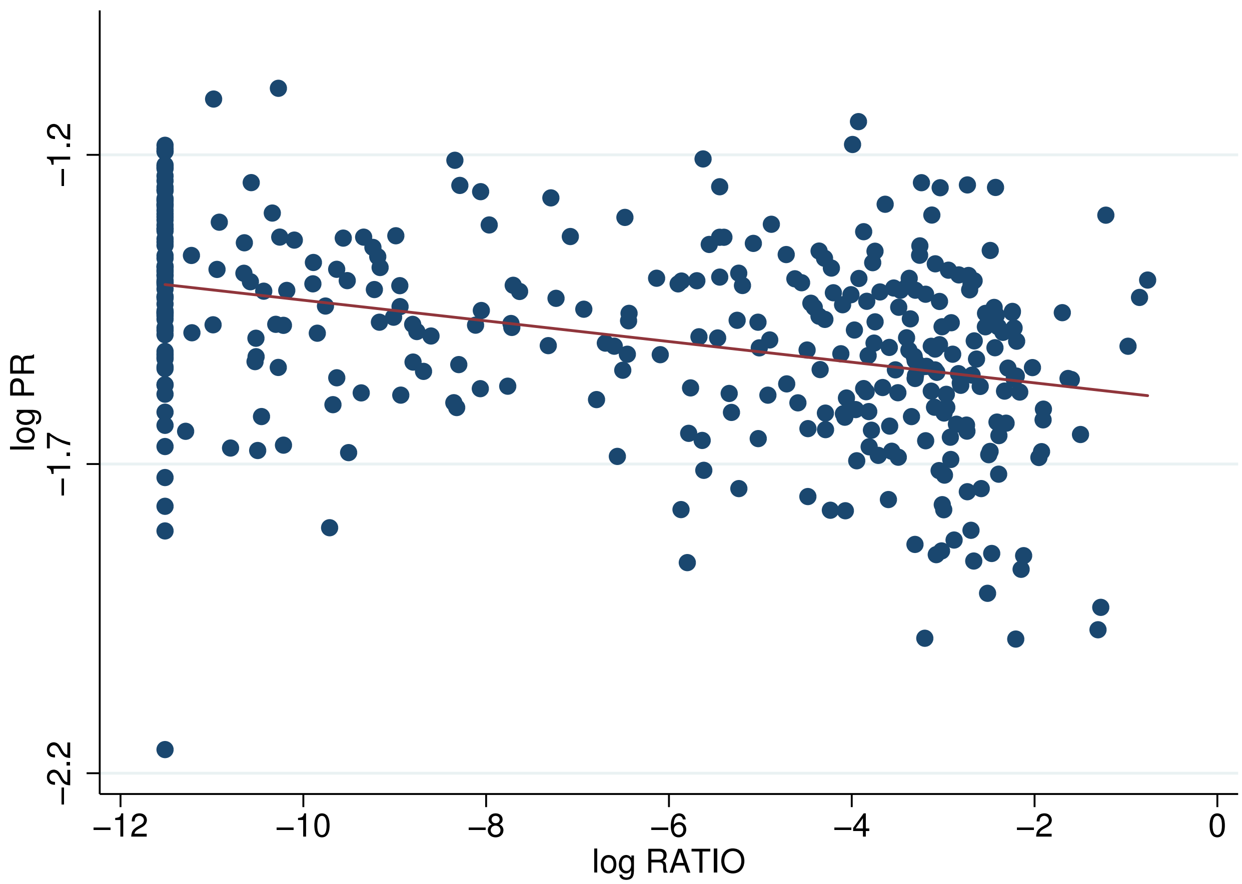

Figure 3: HPIAPP by Income MSA Group and Subprime MSA Group

The figure plots annual home price appreciation (HPIAPP) by the income MSA group and by subprime MSA group. HPIAPP is the average home price appreciation within group by year. We use the home price index from the Federal Housing Finance Agency.