When Real Estate is the Only Game in Town*

Hyun-Soo Choi† Harrison Hong‡ Jeffrey D. Kubik§

Jeffrey P. Thompson¶

First Draft: January 15, 2013

This Draft: February 8, 2015

Abstract

Household stock and real estate investments are known to be locally biased. We establish that households residing in a Metropolitan Statistical Area (MSA) with few publicly traded firms headquartered there are more likely to purchase investment homes nearby and less likely to own stocks. However, MSAs with few firms might simply have more unsophisticated investors, who are familiar with homes as opposed to stocks. Indeed, households living in low FICO (subprime) zip codes, controlling for wealth and income, are more likely to invest in homes than stocks than households living in high FICO zip codes in the same MSA. But, this difference is smaller for MSAs with few firms, consistent with our only-game-in-town effect. Our effect explains the recent boom and bust cycles for large MSAs like Phoenix and Las Vegas, which had been previously difficult to rationalize using housing supply elasticity alone.

*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, National University of Singapore, the CICF 2013 conference, the NBER Behavioral Finance meeting, KAIST, Yonsei, SKKU, the ABFER Annual conference 2014 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 of the Federal Reserve or the Board of Governors.

†Singapore Management University and SKBI (e-mail: hschoi@smu.edu.sg)

‡Princeton University, NBER and CAFR (e-mail: hhong@princeton.edu)

§Syracuse University (e-mail: jdkubik@maxwell.syr.edu)

¶Federal Reserve Board (e-mail: jeffrey.p.thompson@frb.gov)

1. Introduction

According to the National Association of Realtors (NAR) Home Buyers Survey, investment homes, defined as homes bought for investment as opposed to occupation by the owner, represented on average about 22% of the residential sales market or around one million homes annually between 2003 and 2013. 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. In contrast to vacation homes which are bought for leisure, households characterize these purchases in the same way that they describe their purchases of IBM or Microsoft.

Beyond being a sizeable part of the real estate market, and as can be seen in Panel A of Table 1, demand for these homes seems to be correlated with the recent housing cycle of 2002-2007. In 2003, primary residences accounted for 67% of the market, vacation properties 12% and investment homes 22%. During the peak real estate bubble year of 2005, investment homes rose to 28% of the market while primary residences dropped to 60%. Several interesting studies of this recent housing cycle, including Chinco and Mayer (2012), Li and Gao (2013), and Haughwout et al. (2011), show in various ways that cities with more speculation or investment home purchases during this period experienced the biggest boom and bust pattern. These studies have not, however, studied the determinants of households' investment home demand.

As such, we bring to bear new micro-data and new micro-evidence in support of a speculative motive behind investment home purchases.[1] Our hypothesis is simply that in MSAs where there are few stocks headquartered there, households are more likely to purchase an

[1]The focus of most research in the last twenty years in real estate economics has been on owner-occupied homes or primary residences. 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)). 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)).

investment home nearby. We build on a key fact about household investment behavior: households' investments 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., French and Poterba (1991), 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 firsthand.

Not surprisingly, investment home purchases are also locally biased. According to the same annual NAR national survey of home purchases, Panel B of Table 1 reports the distribution of the distances of the second home purchases from the primary residence of the buyer. The median distance of the investment home from the buyers' primary residence in 2005 was 10 miles. In contrast, the median distance of vacation homes is 220 miles.[2] These numbers are very stable across the many years over which their surveys were conducted. Our idea then is that when there are few stocks headquartered nearby, households naturally turn to substitutes in the form of investment homes nearby.

To test this hypothesis, we turn to data on household investment portfolios from the Federal Reserve Board's Survey of Consumer Finances (SCF). The SCF 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 also have a host of demographic information to use in our analysis.

Importantly, we can measure both the investment home purchases as well the stock investments of these households. We calculate for each SCF household three dummy variables: whether or not they HAVE INVESTMENT REAL ESTATE, HAVE VACATION HOME and HAVE DIRECTLY-HELD STOCKS. In addition, we can also calculate the value of

[2]Note that strong local bias of investment homes are not inconsistent with the well-known findings in Chinco and Mayer (2012) that out of town buyers drove real estate markets like Phoenix and Las Vegas. Out of town buyers need not be buying investment homes and might be moving to these cities instead. Their study examines price implications and do not have household portfolio data. The fractions of out of town buyers are in any event typically small fraction of total purchases.

its investment homes and vacations 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 HOMES IN TOTAL ASSETS, and % DIRECTLY-HELD STOCKS IN TOTAL ASSETS).

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 value of firms headquartered in an MSA to the income in that MSA. They show that RATIO at the state level is inversely related to the price-to-book of companies in that state. Most of the variation in their RATIO variable comes from the book value of firms headquartered in a state. That is, states with fewer companies have higher prices for their stocks due to an only-game-in-town effect. In particular, they find that RATIO is highly persistent and unassociated with the economic growth prospects in area. Their focus is on RATIO at the state level; in our empirical designs, given even closer proximity of real estate investments, we focus our analysis at the MSA level.

We first run logistic regressions of our dependent variables of interest HAVE INVESTMENT REAL ESTATE and HAVE DIRECTLY-HELD STOCKS on log RATIO.[3] About 13% of the households own an investment home and 21% have directly held stock. These regressions control for a host of household characteristics such as household net worth, income, whether or not the household is an unattached female, family size, age, race, education, family structure, and year effects. The controls work as one would expect. For instance, greater income or net worth increase the propensity to own an investment home, while being an unattached female decreases this propensity.

Holding fixed these household characteristics, we find that moving one standard deviation up log RATIO lowers the propensity to own an investment home by around 1%, which is around 7% of the unconditional probability of having an investment home. This one standard

[3]More precisely, we take the log(.00001+RATIO) to account for MSAs that have no stocks located nearby. Our baseline sample focuses on large MSAs with populations greater than 750,000 people so instances of zero RATIO are rare. The results for all MSAs are similar.

deviation move also increases the propensity to have directly held stock by 1.8%, or around 9% of the unconditional probability of having any stocks. These effects are statistically significant and economically significant. They are comparable in size to about half of being an unattached female and about half the size of household income.

We then run tobit regressions with % INVESTMENT REAL ESTATE IN TOTAL ASSETS and % DIRECTLY-HELD STOCKS IN TOTAL ASSETS as the dependent variables of interest. The mean percent of households assets in investment real estate is around 3% and the mean percent in directly held stocks is 2%. Being in a high stock supply MSA lowers the percent of total assets in investment real estate by 1.56% or nearly 60% of the base and increases the percent of total assets in stocks by 1.56% or nearly 75% of the base. The logits and tobits suggest that there is both an increase in the discrete decision of whether or not to buy an investment home as well as an increase in the dollar value exposed to investment real estate.

Our main worry in interpreting these regressions is that of geographic heterogeneity in investor sentiment. The east and west coasts of the United States have more stocks headquartered there than in the other more remote regions of the country. Since Hong et al. (2004) find that Bible Belt States have abnormally low stock ownership controlling for household wealth and education, one alternative is that the Census Divisions where more companies locate have a more urban population with a sentiment for stocks. So log RATIO might simply be picking up unobserved regional variation in investor sophistication (i.e. unsophisticated investors in the Bible Belt are more familiar with real estate as opposed to stocks) as opposed to our only-game-in-town effect.

To deal with this concern regarding heterogeneity in investor sophistication, we first add Year x Census Division Effects into our regression specification, which then measures our effect using only MSA variation within the same Census Division. Note that the Census Division Effects are allowed to vary by year. So we only use MSA variation within the Bible Belt or within the west coast within a year to measure our effect but make no comparisons

between MSAs in the Bible Belt to MSAs in the west coast. Our point estimates are little changed.

Of course, investor sophistication need not differ only across Census Divisions. MSAs with few local firms might simply have more unsophisticated investors, who are familiar with homes as opposed to stocks. Indeed, we show that this alternative is a plausible one. We find that less educated, non-white households, controlling for wealth and income, are more likely to invest in homes rather than stocks. Moreover, households living in low FICO (subprime) zip codes, are controlling for wealth and income, more likely to invest in homes than stocks than households living in high FICO zip codes in the same MSA. These results echo those in Mian and Sufi (2009) who also use zip code level FICO scores to measure investor sophistication.

To address this concern, we use a difference-in-differences strategy. Because there are many households and zip codes within an MSA, we can use within MSA FICO score variation to measure our effect. In other words, we can have Year x MSA Effects when we interact log RATIO with FICO. We find that the difference between investment home outcomes for those in low versus high FICO zip codes is smaller for MSAs with few firms, consistent with our only-game-in-town effect. In other words, in MSAs with few local firms, the behaviors of unsophisticated and sophisticated are much more similar than in MSAs with many local firms. That is, everyone, even sophisticated households who are open to homes and stocks, is tilted toward investment homes in MSAs with few local firms.

We then address a number of other more minor concerns regarding heterogeneity in the economic condition of MSAs. High RATIO areas with lots of companies are desirable and expensive. So we simply see less investment homeownership to the extent homes are just more expensive compared to stocks. This would bias upward our estimates. High RATIO MSAs with more companies around are better served by the banking sector and households end up with easier access to credit. This heterogeneity would actually be biasing down our estimates since we would see more investment homeownership in high RATIO areas as a

result of easier access to credit.

To deal with these concerns, we follow Hong et al. (2008) in adding time-varying MSA characteristics such as housing affordability, unemployment rate and past home price appreciation as controls in our baseline regressions. The point estimates are little changed. In addition, we show that there are no effects for vacation homes, which are used primarily for consumption and which are located far away from the owners. This is a comforting check of the premise of our empirical design.

The SCF data does not identify the distance of the investment home from the investor's primary residence. Given that the vast majority of investment home purchases are within 30 miles, it is probably safe to assume that this figure is representative of the SCF sample as well since the SCF is meant to be a representative sample similar to the NAR surveys. Nonetheless, it would be comforting to verify that this is indeed the case. To do so, we exploit a special survey within the SCF survey which was conducted after the 2007 wave. This special survey reinterviewed the respondents of the 2007 survey. As a result, we can track the changes in the household portfolios for this 2007 wave and relate these to changes in local home prices and local stock prices. We find, consistent with the well known local bias results, the real estate investment values and stock investment values closely track the those of the local home and stock prices.

Finally, we show that our only-game-in-town effect has housing price implications. Our sample period overlaps with the historic run-up of real estate between 2000 and 2006 and its subsequent crash after 2007 until 2010. Glaeser et al. (2008) show that home prices went through a bigger boom-bust pattern in low housing supply elasticity states, such as New York and California, where land is limited and zoning and development rules are stricter. The rationale is that in low housing supply elasticity states, supply could not quickly adjust and so prices had to. In contrast, in high housing supply elasticity states, price responses were more muted as many more homes were built.

We show that RATIO and housing supply elasticity, as measured by the Saiz (2010), are

not very correlated and they have comparable explanatory power for the boom and bust cycles. Our RATIO measure picks up a significant amount of variation for cities such as Phoenix and Vegas which had very dramatic boom-bust patterns but were not very difficult to build. Both Phoenix and Las Vegas have very low supply of stocks headquartered nearby in contrast to other comparably large MSAs.

Our paper is related to recent work on the speculative motive behind home purchases and home improvements including Choi et al. (2014) and Gyourko and Saiz (2004), and the "homes as stocks" view, in which behavioral triggers such as optimistic beliefs played a key role in the recent housing bubble, advocated in Shiller (2005), Glaeser et al. (2013), Greenspan (2010) among others.

Our paper proceeds as follows. The data and definitions are presented in Section II. The empirical results are presented in Section III. We conclude in Section IV.

2. Data and Definitions

As a survey of household finances and wealth, the Survey of Consumer Finance (SCF) includes some assets that are broadly shared across the population (bank savings accounts) as well as 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).[4] (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.[5]

The SCF joins the observations from the AP and list sample through weighting.[6] 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.

2.1. Merging RATIO with SCF

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[7] to the income in that MSA as in Hong et al. (2008). We use COMPUSTAT for the total book value of firms headquartered in a MSA and the Bureau of Economic Analysis for the income in a MSA. 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 the RATIO, 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.

Figure 1 shows the geographic distribution of RATIO across the US. There is wide vari-

[4]See Wilson and William J. Smith (1983) and Internal Revenue Service 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.

[5]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.

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

[7]For our analysis, we used the 2009 definition of MSA from the Office of Management and Budget Bulletins.

ation across MSA in their RATIO values. Within every region of the country, there are high and low RATIO MSAs. However, there are some geographic patterns. RATIO tends to be higher in the Northeast, for example, where lots of firms are headquartered. The West tends to have many of the lowest RATIO MSAs. In our empirical work, we will make sure that our results are not being solely driven by comparisons of MSAs across regions of the US.

RATIO is available at the MSA-level and is merged into the 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 SCF data by MSA and year.[8]

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. 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.

For our analysis, we focus on MSAs with larger than 750,000 population. Ultimately, there are around 18,000 households 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 91,421 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.

[8]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. There is an internal version of the survey, however, that is available only to the Board of Governors economists working on the SCF that has MSA-level (and lower levels of geography including zip-code and county) identification of where the household resides. These geographic variables were used to merge the RATIO variable into the SCF.

2.2. Definitions of Investment Real Estate and Vacation Homes

A group of dependent variables are represented as shares—they sum the dollar value of the households portion of all investment real estate, vacation home, or directly-held stocks, and divide by total 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.).[9]

% INVESTMENT REAL ESTATE IN TOTAL ASSETS is then the dollar value of investment real estates divided by the household's total 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 STOCKS 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 SCF for 1995, 1998, 2001, 2004,

[9]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.

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 0.10. % 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 otherwise. 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 otherwise. The mean is 0.06 with a standard deviation of 0.24. 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 0.41. 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 value of firms headquartered in a MSA to the income in that MSA. The mean of Log RATIO is -1.13 with a standard deviation of 1.08. MSA Unemployment Rate is unemployment rate from the Bureau of Labor Statistics. The mean is 5.65 and standard deviation is 2.26. MSA Home Price Index is the Federal Housing Finance Agency (FHFA) Housing Price Index at the MSA level has a mean of 163 with a standard deviation of about 54. MSA Past Price Appreciation is the past home price appreciation of the MSA over a year using the FHFA House Price Index, where $\frac{HPI_t - HPI_{t-1}}{HPI_{t-1}}$. It has a mean of 0.03 and a standard deviation of 0.06. MSA Housing Affordability Index is

the Housing Affordability Index from the National Association of Realtors.[10] 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. Since the index is available only for 2009 to 2011, we backfilled the data using time-series average of available Housing Affordability Index. The Housing Affordability Index has a mean of 120 with a standard deviation of about 58. MSA Housing Supply 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 1.46 with a standard deviation of 0.76.

FAMILY SIZE is the number of people in each Primary Economic Unit. Our households have around 2.44 members with a standard deviation of 1.42. Log HOUSEHOLD INCOME has a mean of 10.87, which is around 53 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. HOUSEHOLD NET WORTH has a mean of about 600,000 dollars with a standard deviation of 4,100,000. Table 2 also reports the breakdown of AGE of the head of each household. The sample is meant to be nationally representative and as such we see a distribution across the age cohorts. We also break down RACE. 68% of the households are white. 16% are black. Hispanics account for 10% of the households and Asians or others account for the remaining 5%.

In Table 2, we also break down EDUCATION of the head of each household. 13% have less than a high school education. 28% have high school or GED equivalent. 25% went to some college but do not have a Bachelor's degree. 21% have a Bachelor's degree. 8% have a Master's degree and 6% have an advanced degree such as PhD, JD, MD, or MBA. FAMILY STRUCTURE breaks down situations like whether or not the couple is married and has children. Here LWP stands for the"living with partner".

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

3. Empirical Findings

3.1. RATIO and Investment Home Ownership

We begin in Table 3 by reporting the household-level panel logit regression results, measuring how log RATIO affects dummy variables for investment real estate and stock ownership. In our baseline specification, we estimate a logit model of the asset holding indicator on log RATIO with controls for the following household characteristics: log HOUSEHOLD INCOME, log HOUSEHOLD NET WORTH. UNATTACHED FEMALE, AGE, EDUCATION and RACE. Also, included in all the logit specifications are year effects and a series of other household level controls for FAMILY STRUCTURE, and a linear and square term of FAMILY SIZE. For brevity, we do not report these coefficients on FAMILY STRUCTURE, and a linear and square term of FAMILY SIZE.

In column (1), the dependent variable is the dummy variable HAVE INVESTMENT REAL ESTATE. The coefficient in front of log RATIO is -0.0831 with a t-statistic of -2.11. Consistent with our only-game-in-town hypothesis, households are less likely to own an investment real estate if they live in a high RATIO MSA. The marginal effect is -0.0082. This suggests that a one standard deviation increase in log RATIO (1.08) decreases the probability of a household owning investment real estate by about 0.0082. Given that the unconditional probability is around 0.13, this is about a 7% decline in the probability.

The coefficients on the other control variables of the specification have the expected sign. Higher log HOUSEHOLD INCOME, HOUSEHOLD NET WORTH, and age increase the probability of having an investment real estate. Having lower educational attainment and being non-white are also associated with higher probabilities of owning investment real estate. This is consistent with the idea that more "sophisticated" investors are more likely to invest stocks. Households with less experience with the financial industry are less likely to invest in stocks and therefore are more likely to place their wealth in assets like investment real estate.

This relationship between being a less sophisticated investor and the likelihood of investing in real estate creates one of the biggest confounders that might invalidate our interpretation of the relationship between log RATIO and real estate investment. As seen in Figure 1, there is some spatial correlation in RATIO across the US. RATIO tends to be higher in MSAs in the Northeast and lower in the West and the South. MSAs in the Midwest are likely to be in the middle of the RATIO distribution. We are worried that there might be more sophisticated investors living in Northeast MSAs than other parts of the country. This would create a negative correlation between the log RATIO of the MSA of the household and the likelihood of owning investment real estate that would be caused by the geographic distribution of investor sophistication and not by the causal influence of log RATIO.

Much of our empirical work will involve trying to rule out this alternative explanation for the relationship between log RATIO and investment behavior. We address this issue in several ways. First, we believe that the rich set of household level controls available in the SCF that we add to the logit specifications capture some of the household characteristics that determine whether they are sophisticated investors or not. The standard demographic controls such as household income, wealth, age, education and race seem like the natural observables to condition on. However, there is no guarantee that these observables are adequate. For example, consider two wealthy households with the same other demographic observables but one lives in New York City and the other lives in the South. It might be that on average the household located in the South came to its high wealth more recently than the household in New York City giving that household less time to learn how to be a sophisticated investor. Such information is not available in the SCF.

To attempt to address such unobservables, we next try to identify the relationship between log RATIO and investment real estate behavior without using cross-region variation in log RATIO. We have shown that there are large differences across regions of the US in log RATIO, but we are worried that those across-region differences are correlated with other factors that determine investment behavior. But if we add Census Division by year effects

to the logit specification, we can identify the coefficient on log RATIO only with variation in MSA log RATIO within the same Census Division.[11] Therefore, comparisons of the behavior of households in Northeast MSAs to households in the South are not being used to calculate the effect of log RATIO on the propensity of households to hold investment real estate.

In column (2), we show the results of the logit specification including these Census Division by year effects. The coefficient on log RATIO in column (2) is very similar in magnitude and statistical significance to the previous results. A one standard deviation increase in log RATIO is suggested to decrease the probability that a household holds investment real estate by about 6% of the unconditional mean, almost the same as in column (1). Therefore, our results are not being driven by long-distance comparisons of MSAs. However, adding these Census Division by year effects will not address our sophisticated investor confounding story if within a Census Division there is still a sizable relationship between the log RATIO of MSAs and the percentage of sophisticated investors within that MSA. So even though it is comforting that adding these controls does not change our estimates, it does not completely rule out alternative stories.

Both of these logit specifications indicate that households in MSAs with a higher log RATIO are less likely to hold investment real estate. Panel A of Figure 2 shows a residual scatter plot of the relationship between investment real estate holding and log RATIO at the MSA/year level. The regression specification is the same as the one shown in column (1) of Table 3. Residualized log RATIO averaged at the MSA/year level is plotted against residualized investment real estate holding propensity again at the MSA/year level. The scatter plot shows how the negative relationship between the two variables is identified.

For the rest of the results shown in Table 3, we examine the relationship between log RATIO and the propensity for households to directly hold stocks. Recall that our hypothesis

[11] There are nine Census Divisions. They are New England (CT, ME, MA, NH, RI, VT), Middle Atlantic (NJ, NY, PA), South Atlantic (DE, FL, GA, MD, NC, SC, VA, WV), East South Central (AL, KY, MS, TN), East North Central (IL, IN, MI, OH, WI), West South Central (AR, LA, OK, TX), West North Central (IA, KS, MN, MO, NE, ND, SD), Mountain (AZ, CO, ID, MT, NV, NM, UT, WY), and Pacific (AK, CA, HI, OR, WA).

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. The specifications in columns (3) and (4) are identical to the first two columns of Table 3 except the outcome variable is instead an indicator for holding stocks.

Consistent with our hypothesis, we find in both of our specifications a positive and statistically significant relationship between log RATIO and the propensity to hold stocks directly. A one standard deviation increase in log RATIO is associated with 9% increase in the probability that a household holds stock relative to the unconditional probability. Removing across-region variation in log RATIO by adding Census Division by year effects in the specification in column (4) again does not change the coefficient on log RATIO.

Also, the coefficients on the household level controls are consistent with the sophisticated investor story we told above. Higher educated, white households are substantially more likely to invest in stocks, conditional on income and wealth, than other households.

One might worry that our findings in columns (3) and (4) of Table 3 are 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 results for stock holding. Also, 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.

Just looking at whether or not a household owns investment real estate or stocks might be too coarse a distinction to be meaningful. Therefore, we turn to examining how log RATIO affects the amount of investment real estate and stocks a household holds measured as a percentage of the household's portfolio devoted to those two assets. Table 4 presents

these results. It is identical to Table 3 in structure except for two changes. In columns (1) and (2), the dependent variable is the percentage of a household's portfolio devoted to investment real estate instead of an indicator for holding investment real estate. Similarly, in columns (3) and (4), the dependent variable is the percentage of a household's portfolio devoted to stocks. Because the dependent variables in Table 4 are bounded between 0 and 1, we estimate the models using tobits instead of logits.

Columns (1) and (2) of Table 4 show that the relationship between log RATIO and the amount of households' portfolios devoted to investment real estate is negative and statistically different from zero across both specifications. The effect of a one standard deviation increase in log RATIO on the percentage of assets in investment real estate ranges from a decrease of 46 to 59% relative to the standard deviation of holdings. As with the logit results, the magnitude of the coefficient on log RATIO is unaffected by the inclusion of Census Division by year effects. Panel B of Figure 2 shows the residualized scatter plot of the relationship between log RATIO at the MSA/year level and the residualized average percent holding of investment real estate again at the MSA/year level.

Similarly, in columns (3) and (4), we show there is a positive and statistically significant effect of log RATIO on the proportion of a household's assets devoted to stocks. The tobit coefficients suggest that a one standard deviation increase in log RATIO raises a household's holding of stocks about 72 and 75% relative to the standard deviation of holdings.

In sum, both Tables 3 and 4 show a substantial relationship between the log RATIO of a MSA and the holdings of households in that MSA of investment real estate.

3.2. FICO Diff-in-Diff

As discussed above, we are most worried about the relationship between log RATIO at the MSA level and investor sophistication as an alternative interpretation of our results. We have tried some empirical tests to deal with this concern, but we have not completely ruled out this alternative.

The ideal empirical strategy to address this issue would be to isolate exogenous variation in log RATIO within a MSA over time. For example, if there is a large exogenous increase in log RATIO in a MSA, we could compare the behavior of households before and after this change to see if they decrease their holdings of real estate investments. Assuming the change in log RATIO in that MSA over time is not related to any changes in investor sophistication in that MSA, we would be observing the effect of the change of log RATIO on the change in household behavior that cannot be driven by a relationship between log RATIO and sophistication. However, we have not been able to isolate large changes in MSA log RATIO over time that we can confidently say are not correlated with other factors that might change household investment decisions.

We can try another diff-in-diff strategy that is similar to this ideal. Because we know the zip code of the primary residence of a household in the SCF, we can match the average FICO score of households that live in that zip code to the household's other information. The credit-worthiness of the people who live around a household might be a good indicator of the household's investment sophistication. A high income and wealth household that lives in a high FICO score zip code probably has had that high income and wealth longer than similar households living in low FICO zip codes, giving the household more time to become a sophisticated investor.

We can then classify households as living in low versus high credit risk zip codes. Households in the top 25% of the FICO score distribution are classified as living in low risk zip codes. Households in the bottom 25% of the FICO score distribution are classified as living in a high risk zip code. We leave out the middle 50% of the FICO distribution to make sure we are comparing households across a wide range of average FICO scores. Therefore, within each MSA we have a sample of households that we classify as sophisticated or not. There will be different proportions of these households across MSAs, but the key is within every MSA we have some sophisticated and unsophisticated investors.

With this classification of households, we ask whether log RATIO has a larger effect on

the portfolio decisions of households located in high FICO score zip codes than low score zip codes. To answer this question, we estimate models where an indicator for having investment real estate is regressed on log RATIO, an indicator for living in a high FICO score zip code, an interaction of log RATIO and the high FICO score zip code indicator and other controls. The coefficient of interest is on the interaction term; it tells us whether log RATIO has a different effect on people with more.

One of the benefits of this diff-in-diff specification is that we can include MSA by year effects to the model and still identify the interaction term. This is because there is variation in the interaction term within MSA by year cells because some households live within low and high risk zip codes within the same MSA at the same time. Therefore, we can difference out MSA-level differences in investor sophistication.

The results of these diff-in-diff models are presented in Table 5. Columns (1) and (2) show the results with the dependent variable an indicator for having investment real estate.[12] In column (1), we present the model without the interaction term to show that we obtain the same negative relationship between log RATIO and the probability of holding investment real estate even for this smaller sample of households living in high and low FICO score zip codes. Notice that the coefficient on living in a high FICO zip code is negative, suggesting that it is picking up investor sophistication.

In column (2), we add the interaction term and MSA by year fixed effects to the regression specification. The interaction term is negative and statistically significant from zero, indicating that the negative sensitivity of investment real estate holding to log RATIO is larger for households in low constraint areas. The magnitude of the interaction coefficient in columns (2) suggests that sensitivity is about four times larger for the households living in high FICO zip codes than the average effect for the entire population.

Columns (3) and (4) show similar results for the tobit models. The basic relationship

[12]We use linear probability models for columns (1) through (2) of Table 5 instead of logits. The interaction terms are easier to interpret in linear models. If we estimate the logits presented earlier as linear probability models instead, we find very similar quantitative and qualitative results.

between log RATIO and real estate investment holdings in column (3) is almost the same as previously shown in Table 4. The interactions term in columns (4) is again very large relative to the base effect of log RATIO shown in column (3). The interaction suggests that all of the effect of log RATIO on investment real estate portfolio decisions is again about four times for households in high FICO zip codes than for the average of the entire population. These large interaction estimates suggest that our main specifications might be underestimating the true causal relationship between log RATIO and real estate investment behavior.

3.3. Other Robustness Checks

Another potential story that confounds our results involves differences in the price of real estate across MSAs. With higher real estate prices, it might be more difficult or less attractive for a household to invest locally in real estate. In this story, if MSAs with higher log RATIO have higher real estate prices, then there would be a negative correlation between log RATIO and real estate holdings even if there was no causal relationship between the two. Given that there are large regional differences in real estate prices, specifications that include Census Division by year effects might deal with this issue. But large differences in real estate prices within a region would not be dealt with.

To deal with this concern, we estimate our previous logit and tobit models including additional time-varying MSA-level controls to account for differences across MSAs in economic conditions and the housing market. These variables are designed to control for differences in the affordability of real estate across MSAs over time. We include five MSA-level controls: the MSA Unemployment Rate, the MSA Home Price Index, Past (Home) Price Appreciation of a MSA, the MSA Housing Affordability Index and a measure of the MSA Housing Supply Elasticity.

The results of logit and tobit estimates of the effect of log RATIO on asset holding behavior including these controls are presented in Table 6. The first two columns are the logit estimates. The coefficients on log RATIO for both the investment real estate holding

and stock holding logits are very similar to before. The precision of the estimates falls somewhat, but the magnitudes of the marginal effects are relatively unchanged. In columns (3) and (4) of Table 6, the equivalent tobit estimates are presented. The story is very similar to the logit estimates. The estimated magnitudes are very similar to before, but the precision of the estimates worsens.

There is one more empirical strategy we have tried that we believe helps us rule out this alternative explanation, but again it is not definitive. We have run the same logit and tobit models as we presented in Tables 3 and 4, but instead of using as a dependent variable a measure of household real estate investment decisions, we use measures of household vacation home holdings. In our framework, we do not expect to see much of a relationship between log RATIO and vacation home holding decisions because such homes are typically consumption decisions and not investment decisions. But in the sophisticated or hickish investor framework, one might expect a relationship. If hickish investors do not have good outlets for investment beyond local real estate, they might instead put their money in things like vacation homes, creating a negative correlation between log RATIO and vacation home holdings.

In Table 7, we present the results showing the relationship between log RATIO and vacation home holdings of households. Columns (1) through (3) show the logit estimates for whether a household owns any vacation home. The coefficients are all wrong-signed compared to our hypothesis, but they are all relatively imprecise and small in magnitude. There are a similar set of results for the tobit estimates for the percentage of assets in vacation homes presented in columns (4) through (6). Again, the coefficient on log RATIO is always wrong-signed.

3.4. Local Bias of Portfolios in SCF Sample

Although we have cited substantial evidence that there is local bias in investment real estate investment decisions, we also want to examine whether we can find evidence of this bias in

the SCF. The SCF does not provide information on the location of a household's real estate investments. But we can infer the location of the households' real estate holdings using the 2007 SCF. Households in that wave of the SCF were re-interviewed two years later. Therefore, we can create a short panel that measures the change in a household's assets between 2007 and 2009.

For each MSA with a population greater than 750,000, we sum the value of investment real estate, vacation homes and stock assets of all the households in that MSA in 2007 and 2009. We then calculate the percentage change of those assets over those years. We then correlate those percentage changes in assets with the local return of those assets over the same time period. This will give us a measure of how local these investment are.

The results of these regressions are reported in Table 8. In column (1), we estimate the localness of households' investment real estate assets. We regress at the MSA level the percentage change in investment real estate assets of a MSA on the return to housing in that MSA over the same time period. This MSA housing return is calculated based on FHFA housing price indices. The coefficient on the MSA housing return is about one and statistically different from zero. Changes in local housing prices strongly predict changes in investment real estate assets, suggesting that a large percentage of these investment real estate investments must be local.

In column (2), we estimate a similar regression specification, except that the dependent variable is a measure of the percentage change in vacation home assets in a MSA instead of investment real estate. The relationship between this measure of vacation home holdings and local housing returns is much smaller. The coefficient on MSA housing returns is about 0.27 and statistically insignificant. Like previous work, we find that vacation home investments is substantially less local than investment real estate.

Finally, we use a similar methodology to measure the localness of stock investments of households in the 2007 SCF. In column (3), we regress the percentage change of stock holdings within a MSA on a measure of the stock returns of local stocks over the same time period.

The coefficient on the MSA local stock return is 0.47 with a t-statistic of 1.53. Therefore, our results suggest that the localness of stock investment is not as large as investment real estate but somewhat larger than vacation homes.

3.5. Pricing Implications

Having established that households living in MSAs with less local stocks are more likely to buy investment homes, we then see if there are any pricing implications from an only-game-in-town effect for real estate. In particular, our analysis here is motivated by the interest in understanding the fluctuations in home prices between the years of 2002 to 2010.

Two main ingredients have been advocated for the recent cycle. The first is easier credit (see, e.g., Mian and Sufi (2009)) whereby it was easier for Americans to get credit to buy homes than at any time before and released pent-up demand for homes due to financial constraints. The second is speculation or optimistic beliefs that home prices would keep rising (see, e.g.m Shiller (2005), Glaeser et al. (2013), Greenspan (2010)). These forces both operate at the national level.

Glaeser et al. (2008) point to an important mediating force across cities, which is housing supply elasticity. They show that home prices went through a bigger boom-bust pattern in low housing supply elasticity states, such as New York and California, where land is limited and zoning and development rules are stricter and as measured by the Saiz (2010) measure.

Our analysis suggests a second important mediating force across cities, which is the supply of local stocks. To see this, consider Figure 3. We sort the MSAs into roughly 3 housing supply elasticity groups (Lowest Elasticity 1/3, Middle Elasticity 1/3 and Highest Elasticity 1/3). Then independently, we sort the MSAs into a low versus high RATIO group. The low RATIO group is in Blue and the high RATIO group in Red.

In Panel A, we report a rolling 4 quarter average of home price appreciation across six groups over the period of 1997 to 2010. First, notice that in the highest housing supply elasticity states, there is not a significant different between the low versus high RATIO

groups. The blue low RATIO line only peaks above the red line ever so slightly in the 2006 and 2007 years. This is consistent with Glaeser et al. (2008) who show that housing supply elasticity ought to dampen down any effects of speculative demand on prices.

Next consider the lowest housing supply elasticity MSAs. We see here that there are now significant deviations between the blue low RATIO MSAs and the red high RATIO MSAs. There is more of a cycle for the MSAs where real estate is the only game in town. The two blue and red lines track each other well from 1997 to 2000 and then the blue line rises significantly faster and above the red line, only then to crash below it after 2007. One interpretation of this finding is that low RATIO states have more speculative investment home demand than high RATIO states. But financial constraints kept this differential pent-up. But it gets released with the easing of credit in the early 2000s and the low RATIO MSAs end up with a more of a cycle.

Now consider the middle elasticity MSAs. The middle elasticity MSAs are of interest because cities like Phoenix and Las Vegas, which had the biggest housing cycles, are in this middle elasticity MSA group. In other words, the Saiz (2010) measure does not capture this variation as a number of authors have pointed out. It turns our that our only-game-in-town measure does pick up this variation as can be seen from Panel A. The blue low RATIO group, which includes Phoenix and Las Vegas, experience a dramatic boom and bust cycle. We see that the blue line peaks much above the red line in the 2005 and 2006 years and falls much below the red line in the 2008 to 2010 period.

In Panel B, we use another metric to make the same point, which is the percentage change in the MSA price-to-rent ratio. We can think of the price-to-rent ratio as an over-valuation indicator and we can see how it varies from year to year. One gets a very similar set of conclusions from this metric as from the price appreciation metric in Panel A.

In Figure 4, we plot the cumulative dollar return to investing in the low versus high RATIO housing portfolios across the three groups of lowest, middle and highest housing supply elasticity groups. The results are in line with Figure 3. There are more dramatic

run-ups in prices for the low RATIO compared to the high RATIO areas for the lowest and middle elasticity groups during the period of 2000-2006. There are also more dramatic crashes for the low RATIO compared to the high RATIO areas during the period of 2007-2010.

In Table 9, we present the results from Figures 3 and 4 using regressions in which we regress the various measures of booms (busts) during the period of 2000 to 2006 (2007 to 2010) on log RATIO and housing supply elasticity. We also add in other MSA covariates. We see that log RATIO actually has comparable explanatory power to housing supply elasticity for the housing cycles across cities.

4. Conclusion

Using data from the Survey of Consumer Finances, we establish that households residing in a Metropolitan Statistical Area (MSA) with few publicly traded firms headquartered there are more likely to purchase an investment home nearby. Households in these areas are also less likely to own stocks. We use a difference-in-differences estimation strategy to show that we are capturing demand effects as opposed to credit supply differences across MSAs. Using FICO score variations within an MSA, we find that our only-game-in-town effect is larger for more credit-worthy households. Our only-game-in-town effect has comparable explanatory power to the well-known housing supply elasticity measure for explaining the cross-city variation in the intensity of the recent housing cycle in the United States.

References

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

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. 2014. Speculating on Home Improvements. *Journal of Financial Economics* 111:609–624.

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

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* 64:198–217.

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* 56:1053–1073.

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

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* .

Hong, H., J. D. Kubik, and J. C. Stein. 2004. Social Interaction and Stock-Market Participation. *Journal of Finance* 59:137–163.

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* 90:20–37.

Huberman, G. 2001. Familiarity Breeds Investment. *Review of Financial Studies* 14:659–680.

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* .

Kennickell, A. B. 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* .

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* 124:1449–1496.

Saiz, A. 2010. The Geographic Determinants of Housing Supply. *Quarterly Journal of Economics* 125:1253–1296.

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* 120:763–789.

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

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: Characteristics of Investment and Vacation Home Purchases

Panel A reports the number of home sales by intended use from the *Investment and Vacation Home Buyers Survey* by the National Association of Realtors. The share in total home sales are reported in the parenthesis. Panel B reports the distance of second homes from primary residence. The number of respondents for the Survey was 8205 for year 2004, 1034 for year 2005, 1965 for year 2007, 1924 for year 2008, 1930 for year 2009, 1895 for year 2010, 2241 for year 2011, 2326 for year 2012, and 2008 for year 2013.

Panel A : *Home Sales by Intended Use*

	Primary Residences	Vacation Properties	Investment Properties
2003	4,844,000 (67%)	850,000 (12%)	1,572,000 (22%)
2004	5,106,000 (64%)	872,000 (11%)	2,003,000 (25%)
2005	5,023,000 (60%)	1,019,000 (12%)	2,317,000 (28%)
2006	4,817,000 (64%)	1,067,000 (14%)	1,646,000 (22%)
2007	3,925,000 (67%)	670,000 (12%)	1,221,000 (21%)
2008	3,207,000 (70%)	436,000 (9%)	951,000 (21%)
2009	3,441,000 (73%)	471,000 (10%)	801,000 (17%)
2010	3,294,000 (73%)	469,000 (10%)	749,000 (17%)
2011	2,785,000 (61%)	502,000 (11%)	1,233,000 (27%)
2012	3,268,000 (65%)	553,000 (11%)	1,207,000 (24%)
2013	3,697,000 (67%)	717,000 (13%)	1,104,000 (20%)

Panel B : *Distance from Primary Residence*

Vacation Properties (%)	2013	2012	2011	2010	2009	2008	2007	2005	2004
5 miles or less	2	7	2	2	2	2	4	3	
6 to 10 miles	4	4	3	3	4	2			
11 to 15 miles	6	2	2	3	1	1	2		44
16 to 20 miles	6	5	4	4	3	4	4	4	
21 to 50 miles	13	6	10	8	8	7	7	8	19
51 to 100 miles	15	10	14	11	16	19	18	19	
101 to 500 miles	20	20	27	27	27	30	30	31	23
501 to 1000 miles	12	15	11	11	12	14	12	9	3
1001 miles or more	22	31	26	30	28	22	22	25	10
Median (miles)	180	435	305	375	348	316	287	220	49

Investment Properties (%)	2013	2012	2011	2010	2009	2008	2007	2005	2004
5 miles or less	15	24	17	20	18	24	16	52	
6 to 10 miles	15	13	9	13	12	11	11		
11 to 15 miles	10	6	12	8	9	8	8	17	69
16 to 20 miles	10	7	10	15	9	11	13		
21 to 50 miles	12	15	13	12	11	8	15	6	9
51 to 100 miles	9	6	10	4	9	7	10	5	
101 to 500 miles	10	10	10	10	11	12	12	8	5
501 to 1000 miles	8	8	7	5	7	8	6	3	3
1001 miles or more	12	11	13	13	14	11	10	9	14
Median (miles)	20	21	25	19	24	19	27	10	18

Table 2: Summary Statistics

We report summary statistics of the Survey of Consumer Finance for 1995, 1998, 2001, 2004, 2007, and 2010 waves. The sample consists of households in MSAs with a population greater than 750,000. HAVE INVESTMENT REAL ESTATE is an indicator that the family owns investment real estate. HAVE VACATION HOME is an indicator that a family owns a vacation home. HAVE DIRECTLY-HELD STOCKS is an indicator that the family directly owns stock. % INVESTMENT REAL ESTATE IN TOTAL ASSETS is the share of a family's assets invested in investment real estate. % VACATION HOME IN TOTAL ASSETS is the share the family's assets invested in a vacation home. % DIRECTLY-HELD STOCKS IN TOTAL ASSETS is the share the family's assets invested directly in stock. 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. MSA Unemployment Rate is unemployment rate from the Bureau of Labor Statistics. MSA Housing Price Index is the Federal Housing Finance Agency (FHFA) Housing Price Index at the MSA level. MSA Past Price Appreciation is the annual past home price appreciation in an MSA using the FHFA House Price Index, where $(HPI_t - HPI_{t-1})/HPI_{t-1}$. MSA Housing Affordability Index 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. MSA Housing Supply Elasticity is the housing supply elasticity by MSA from Saiz (2010). FAMILY SIZE is the number of people in each Primary Economic Unit. log HOUSEHOLD INCOME is the log of household income. HOUSEHOLD NET WORTH is the household's net worth. UNATTACHED FEMALE is an indicator that the family is headed by an unattached woman. AGE, RACE, EDUCATION, and FAMILY STRUCTURE are also reported. LWP stands for the "living with partner" in the FAMILY STRUCTURE variable.

Variables		Obs	Mean	Std.	Max	Min
HAVE INVESTMENT REAL ESTATE		91421	0.13	0.34	1	0
HAVE VACATION HOME		91421	0.06	0.24	1	0
HAVE DIRECTLY-HELD STOCKS		91421	0.21	0.41	1	0
% INVESTMENT REAL ESTATE IN TOTAL ASSETS		91421	0.03	0.10	1.08	0
% VACATION HOME IN TOTAL ASSETS		91421	0.01	0.04	0.87	0
% DIRECTLY-HELD STOCKS IN TOTAL ASSETS		91421	0.02	0.08	1	0
log RATIO		91421	-1.13	1.08	1.11	-11.51
MSA Unemployment Rate		91421	5.65	2.26	15.87	2.14
MSA Home Price Index		91421	162.91	53.69	323.15	100.75
MSA Past Price Appreciation		91421	0.03	0.06	0.23	-0.17
MSA Housing Affordability Index		91421	119.51	57.69	383.4	51.17
MSA Housing Supply Elasticity		91421	1.46	0.76	4.00	0.63
FAMILY SIZE		91421	2.44	1.42	13	1
log HOUSEHOLD INCOME		91421	10.87	1.14	19.64	0
HOUSEHOLD NET WORTH		91421	602109	4090744	4.80E+09	-2.74E+07
UNATTACHED FEMALE		91421	0.27	0.45	1	0
AGE	<35	91421	0.23	0.42	1	0
	35 – 49	91421	0.33	0.47	1	0
	50 – 64	91421	0.24	0.43	1	0
	≥65	91421	0.19	0.40	1	0
RACE	White	91421	0.68	0.47	1	0
	Black	91421	0.16	0.36	1	0
	Hispanic	91421	0.10	0.30	1	0
	Asian & Other	91421	0.05	0.21	1	0
EDUCATION	Less than High School	91421	0.13	0.33	1	0
	High School or GED only	91421	0.28	0.45	1	0
	Some college (or Associate's), but no Bachelor's	91421	0.25	0.43	1	0
	Bachelor's	91421	0.21	0.41	1	0
	MA or MS(non-MBA), plus nursing degree	91421	0.08	0.26	1	0
	Advanced (PhD, JD, MBA, MD)	91421	0.06	0.24	1	0
FAMILY STRUCTURE	Not Married/LWP+Children	91421	0.13	0.33	1	0
	Not Married/LWP+no Children+Head(<55)	91421	0.16	0.37	1	0
	Not Married/LWP+no Children+Head(≥55)	91421	0.13	0.34	1	0
	Married/LWP+Children	91421	0.32	0.47	1	0
	Married/LWP+no Children	91421	0.26	0.44	1	0

Table 3: Investment Real Estate Ownership, Directly-held Stocks Ownership, and log RATIO

We report logit estimates of the relationship between log RATIO and the likelihood that a household owns investment real estate and stocks. The sample consists of SCF households in MSAs with population greater than 750,000. The dependent variable in columns (1) through (2) is an indicator that the family owns investment real estate. The dependent variable in columns (3) through (4) is an indicator that the family directly owns stock. The independent variables include log RATIO, 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) and log HOUSEHOLD INCOME, the usual income of the family and log HOUSEHOLD NET WORTH, the wealth of family (not including real estate in columns (1)-(2) and not including stock holdings in columns (3)-(4)). Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log. UNATTACHED FEMALE is an indicator that the family is headed by an unattached woman. Age, race, and education-level of the head of the family are also reported. For brevity, family structure, and a linear and square term of the number of people in the family are not reported. The table reports point estimates with t-statistics in parentheses. All the standard errors are clustered at the MSA level. ***, **, * denotes 1%, 5%, and 10% statistical significance. The average marginal effect is in brackets. Economic Significance is the percentage change in the dependent variable implied by a one standard deviation increase in log RATIO.

		(1)	(2)	(4)	(5)
		HAVE INVESTMENT		HAVE DIRECTLY-HELD	
Variables		REAL ESTATE		STOCK	
log RATIO		-0.0831**	-0.0754***	0.131***	0.136***
		(-2.11)	(-2.67)	(4.09)	(3.82)
		[-.0082]	[-.0073]	[.0172]	[.0176]
Economic Significance		-0.07	-0.06	0.09	0.09
log HOUSEHOLD INCOME		0.21***	0.211***	0.357***	0.359***
		(4.47)	(4.72)	(6.83)	(6.43)
log HOUSEHOLD NET WORTH		0.576***	0.586***	0.436***	0.436***
		(13.71)	(14.15)	(13.09)	(13.63)
UNATTACHED FEMALE		-0.275**	-0.275**	-0.216***	-0.217***
		(-2.25)	(-2.20)	(-3.18)	(-3.14)
Age	Age 35-49	0.278**	0.282*	-0.225**	-0.22**
		(2.28)	(2.37)	(-2.12)	(-2.06)
	Age 50-64	0.536***	0.561***	-0.346***	-0.333***
		(4.22)	(4.56)	(3.17)	(-3.08)
	Age 65+	0.469***	0.497***	-0.0794	-0.0735
		(3.4)	(3.77)	(-0.74)	(-0.71)
Race	Black	0.408***	0.416***	-0.545***	-0.54***
		(3.89)	(4.08)	(-8.06)	(-7.86)
	Hispanic	0.237	0.203	-1.004***	-0.981***
		(1.6)	(1.25)	(-6.24)	(-6.25)
	Asian and Other	0.137	0.161	0.0736	0.0794
		(1.03)	(1.2)	(0.79)	(0.81)
Education	High School Degree	-0.132	-0.145	0.481***	0.488***
		(-0.93)	(-1.00)	(3.54)	(3.59)
	Some College	-0.009	-0.021	0.812***	0.836***
		(0.08)	(-0.17)	(5.24)	(5.43)
	Bachelors Degree	-0.092	-0.101	1.186***	1.208***
		(-0.68)	(-0.75)	(7.85)	(7.95)
	MA or MS	-0.0142	-0.023	1.179***	1.201***
		(-0.09)	(-0.15)	(6.94)	(7.11)
	Advanced Degree	-0.255*	-0.266*	1.073***	1.096***
		(-1.76)	(-1.82)	(6.54)	(6.6)
Other Household Controls		Yes	Yes	Yes	Yes
Year Effects		Yes	—	Yes	—
Year x Census Division Effects		No	Yes	No	Yes
Observations		91421	91421	91421	91421

Table 4: % Share of Investment Property, % Share of Directly-held Stocks, and log RATIO

We report tobit estimates of the relationship between log RATIO and the amount that a family owns of investment real estate and stocks. The sample consists of SCF households in MSAs with population greater than 750,000. The dependent variable in columns (1) through (2) is the share of a family's assets invested in investment real estate. The dependent variable in columns (3) through (4) is the share the family's assets invested directly in stock. The independent variables include log RATIO, 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) and log HOUSEHOLD INCOME, the usual income of the family and log HOUSEHOLD NET WORTH, the wealth of family (not including real estate in columns (1)-(2) and not including stock holdings in columns (3)-(4)). Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log. UNATTACHED FEMALE is an indicator that the family is headed by an unattached woman. Age, race, and education-level of the head of the family are also reported. For brevity, family structure, and a linear and square term of the number of people in the family are not reported. The table reports point estimates with t-statistics in parentheses. All the standard errors are clustered at the MSA level. ***, **, * denotes 1%, 5%, and 10% statistical significance. Economic Significance is the percentage change in the standard deviation of dependent variable implied by a one standard deviation increase in log RATIO.

		(1)	(2)	(4)	(5)
		% INVESTMENT REAL ESTATE		% DIRECTLY-HELD STOCKS	
Variables		IN TOTAL ASSETS		IN TOTAL ASSETS	
log RATIO		-0.0156**	-0.0122**	0.0156***	0.015***
		(-2.27)	(-2.24)	(4.71)	(4.1)
Economic Significance		-0.59	-0.46	0.75	0.72
log HOUSEHOLD INCOME		0.0393***	0.0377***	0.045***	0.045***
		(3.85)	(3.77)	(5.65)	(5.75)
log HOUSEHOLD NET WORTH		0.0933***	0.0938***	0.0278***	0.0277***
		(11.58)	(12.04)	(5.99)	(6.11)
UNATTACHED FEMALE		-0.0599**	-0.0584**	-0.03***	-0.0306***
		(-2.3)	(-2.41)	(-3.68)	(-3.74)
Age	Age 35-49	0.0474**	0.0494**	-0.0252***	-0.025***
		(1.94)	(2.1)	(-2.72)	(-2.70)
	Age 50-64	0.0975***	0.102***	-0.0298**	-0.0291**
		(4.06)	(4.42)	(-2.57)	(-2.53)
	Age 65+	0.0893***	0.0933***	0.026**	0.0267**
		(3.18)	(3.41)	(2.45)	(2.57)
Race	Black	0.0898***	0.0916***	-0.0699***	-0.0688***
		(3.97)	(4.14)	(-8.89)	(-8.88)
	Hispanic	0.0497*	0.0393	-0.108***	-0.107***
		(1.68)	(1.19)	(-6.32)	(-6.48)
	Asian and Other	0.0363	0.0368	-0.00229	-0.00132
		(1.29)	(1.27)	(-0.20)	(-0.10)
Education	High School Degree	-0.0401	-0.0408	0.0666***	0.0657***
		(-1.34)	(-1.35)	(4.41)	(4.35)
	Some College	-0.0106	-0.0117	0.104***	0.105***
		(-0.42)	(-0.46)	(6.05)	(6.21)
	Bachelors Degree	-0.0158	-0.0163	0.143***	0.144***
		(-0.55)	(-0.56)	(8.22)	(8.32)
	MA or MS	0.0001	0.0000	0.142***	0.143***
		(0.01)	(0.02)	(7.8)	(8.08)
	Advanced Degree	-0.0385	-0.0383	0.139***	0.139***
		(-1.25)	(-1.23)	(7.51)	(7.51)
Other Household Controls		Yes	Yes	Yes	Yes
Year Effects		Yes	—	Yes	—
Year x Census Division Effects		No	Yes	No	Yes
Observations		91421	91421	91421	91421

Table 5: Difference-in-Differences Estimate: Using within MSA FICO score variation

We report estimates of the relationship of log RATIO on a familys holdings of investment real estate, splitting up the effect of log RATIO for households living in low and high credit risk zip codes. Within MSAs with a population greater than 750,000, the sample consists of families who live in a zip code in the highest and lowest 25% of the credit risk distribution (measured by average FICO scores). The dependent variable in columns (1) and (2) is an indicator for whether the family owns investment real estate. The dependent variable in columns (3) and (4) is the share of investment real estate of a familys assets. Columns (1) and (3) redo the results in Table 3 and 4 to show that main results are still valid in the subsample. The independent variables are the same as in Table 3 except for a couple of differences. High FICO Zip Code is an indicator that the family lives in a low credit risk zip code. That is, its average FICO score is in the highest 25% of the US distribution. log RATIO x High FICO Zip Code is the interaction of that indicator and log RATIO. Year x MSA Effects are a full set of MSA by year dummy interactions. Because of the inclusion of MSA x Year Effects, the coefficient on log RATIO is not identified. The coefficients for the first two columns are from a OLS model. The coefficients for the last two columns are from a tobit model. Other Household Controls include age, race, and education-level of the head of the family, family structure, and a linear and square term of the number of people in the family. The table reports point estimates with t-statistics in parentheses. All the standard errors are clustered at the MSA level. ***, **, * denotes 1%, 5%, and 10% statistical significance.

	(1)	(2)	(3)	(4)
Variables	HAVE INVESTMENT REAL ESTATE		% INVESTMENT REAL ESTATE IN TOTAL ASSETS	
log RATIO x High FICO Zip Code		-0.0375**		-0.0408***
		(-2.53)		(-14.57)
log RATIO	-0.0087	—	-0.0111	—
	(-1.38)		(-1.33)	
High FICO Zip Code	-0.0209	-0.0561**	-0.0348	-0.0444***
	(-1.38)	(-2.44)	(-1.63)	(-11.24)
log HOUSEHOLD INCOME	0.0448***	0.045***	-0.0022	0.0052***
	(7.63)	(7.08)	(-0.24)	(13.8)
log HOUSEHOLD NET WORTH	0.0146***	0.0135***	0.0372***	0.0101***
	(7.09)	(6.14)	(4.28)	(30.7)
UNATTACHED FEMALE	-0.0142	-0.0146	-0.015	0.0061
	(-0.89)	(-0.94)	(-0.68)	(1.19)
Other Household Controls	Yes	Yes	Yes	Yes
Year Effects	Yes	—	Yes	—
Year x MSA Effects	No	Yes	No	Yes
Observations	44550	44550	44550	44550

Table 6: Accounting for Time-Varying MSA Economic Conditions

We report estimates of the relationship between log RATIO and measures whether a family owns of investment real estate and stocks, controlling MSA characteristics. The dependent variable in column (1) is an indicator that the family owns investment real estate. The dependent variable in column (2) is an indicator that the family directly owns stock. The coefficients in columns (1)-(2) are from a logit model. The dependent variable in column (3) is the share of a family's assets invested in investment real estate. The dependent variable in column (4) is the share the family's assets invested directly in stock. The coefficients in columns (3)-(4) are from a tobit model. The independent variables include log RATIO, 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) and log HOUSEHOLD INCOME, the usual income of the family and log HOUSEHOLD NET WORTH, the wealth of family (not including real estate in columns (1) and (3) and not including stock holdings in columns (2) and (4)). Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log. UNATTACHED FEMALE is an indicator that the family is headed by an unattached woman. Other Household Controls include age, race, and education-level of the head of the family, family structure, and a linear and square term of the number of people in the family. The table reports point estimates with t-statistics in parentheses. All the standard errors are clustered at the MSA level. ***, **, * denotes 1%, 5%, and 10% statistical significance. For columns (1)-(2), Economic Significance is the percentage change in the dependent variable implied by a one standard deviation increase in log RATIO. For columns (3)-(4), Economic Significance is the percentage change in the standard deviation of dependent variable implied by a one standard deviation increase in log RATIO.

Variables	(1) HAVE INVESTMENT REAL ESTATE	(2) HAVE DIRECTLY-HELD STOCK	(3) % INVESTMENT REAL ESTATE IN TOTAL ASSETS	(4) % DIRECTLY-HELD STOCKS IN TOTAL ASSETS
log RATIO	-0.069*	0.122***	-0.0117	0.0144***
	(-1.68)	(3.37)	(-1.59)	(4.2)
	[-.0068]	[.0160]		
Economic Significance	-0.05	0.08	-0.44	0.69
log HOUSEHOLD INCOME	0.214***	0.36***	0.0394***	0.0453***
	(4.59)	(6.45)	(3.86)	(5.61)
log HOUSEHOLD NET WORTH	0.583***	0.439***	0.093***	0.0279***
	(13.98)	(13.14)	(11.73)	(5.97)
UNATTACHED FEMALE	-0.275**	-0.216***	-0.0597**	-0.0298***
	(-2.25)	(-3.17)	(-2.31)	(-3.65)
MSA Unemployment Rate	-0.041	-0.011	-0.0041	-0.0009
	(-1.48)	(-0.42)	(-0.81)	(-0.31)
MSA Home Price Index	-0.001	-0.001	0.0001	-0.0001
	(-1.11)	(-0.67)	(0.59)	(-0.80)
MSA Past Price Appreciation	-1.47*	-0.705	-0.148	-0.056
	(-1.89)	(-0.73)	(-0.99)	(-0.54)
MSA Housing Affordability Index	0.046	-0.075	0.0075	-0.0112*
	(0.73)	(-1.27)	(0.55)	(-1.70)
MSA Housing Supply Elasticity	0.436**	0.105	0.084**	0.0138
	(2.41)	(0.61)	(2.39)	(0.67)
Other Household Controls	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Observations	91421	91421	91421	91421

Table 7: Vacation Home Ownership, % Share of Vacation Home, and log RATIO

We report estimates of the relationship between log RATIO and measures whether a family owns of a vacation home. The sample consists of SCF households in MSAs with population greater than 750,000. The dependent variable in columns (1) through (3) is an indicator that a family owns a vacation home. The coefficients in those columns are from a logit model. The average marginal effect is in brackets. The dependent variable in columns (4) through (6) is the share the family's assets invested in a vacation home. The coefficients in those columns are from a tobit model. The independent variables include log RATIO, 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) and log HOUSEHOLD INCOME, the usual income of the family and log HOUSEHOLD NET WORTH, the wealth of family (not including real estate). Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log. UNATTACHED FEMALE is an indicator that the family is headed by an unattached woman. Other Household Controls include age, race, and education-level of the head of the family, family structure, and a linear and square term of the number of people in the family. For columns (3) and (6), several MSA controls are added to the specification. The table reports point estimates with t-statistics in parentheses. All the standard errors are clustered at the MSA level. ***, **, * denotes 1%, 5%, and 10% statistical significance. For columns (1)-(3), Economic Significance is the percentage change in the dependent variable implied by a one standard deviation increase in log RATIO. For columns (4)-(6), Economic Significance is the percentage change in the standard deviation of dependent variable implied by a one standard deviation increase in log RATIO.

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	HAVE VACATION HOME			% VACATION HOME IN TOTAL ASSETS		
log RATIO	0.0398	0.0608*	0.0041	0.0058*	0.0068**	0.0007
	(1.04)	(1.72)	(0.1)	(1.67)	(2.01)	(0.17)
	[.0020]	[.0031]	[.0002]			
Economic Significance	0.06	0.06	0.004	0.98	1.14	0.12
log HOUSEHOLD INCOME	0.271***	0.273***	0.27***	0.0394***	0.0393***	0.0391***
	(5.36)	(5.41)	(5.34)	(3.86)	(5.88)	(5.86)
log HOUSEHOLD NET WORTH	0.43***	0.428***	0.433***	0.0419***	0.0411***	0.0418***
	(12.68)	(12.19)	(12.66)	(9.61)	(9.69)	(9.68)
UNATTACHED FEMALE	0.314	0.313	0.308	0.0291	0.0299	0.0289
	(1.49)	(1.48)	(1.47)	(1.18)	(1.24)	(1.17)
MSA Unemployment Rate			-0.0881**			-0.0096**
			(-2.20)			(-1.97)
MSA Home Price Index			0.0000			0.0001
			(-0.10)			(0.52)
MSA Past Price Appreciation			-0.447			0.0217
			(-0.52)			(0.19)
MSA Housing Affordability Index			-0.0372			-0.0081
			(-0.53)			(-1.05)
MSA Housing Supply Elasticity			-0.0602			-0.0147
			(-0.22)			(-0.45)
Other Household Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	—	Yes	Yes	—	Yes
Year x Census Division Effects	No	Yes	No	No	Yes	No
Observations	91421	91421	91421	91421	91421	91421

Table 8: Measuring Local Bias in the 2007 SCF

We report OLS estimates of the relationship between the percentage change of households holdings of different assets in a MSA over time and the return on housing in that MSA over the same time period. The sample consists of MSAs with a population greater than 750,000 with households in the 2007 SCF. The dependent variable in column (1) is the percentage change of investment home assets between 2007 and 2009 of 2007 SCF households within a MSA. The dependent variable in column (2) is a similar measure using the change in vacation home assets of households in the 2007 SCF. The dependent variable in column (3) is the same percentage change but for stocks. The independent variables are a constant and a measure of the return to housing in a MSA between 2007 and 2009 using FHFA house price index in the first two columns. In column (3), the independent variable of interest is the return to local stocks over that time period. All regressions are weighted by the amount of the asset held in the MSA in 2007. t-statistics are in parentheses. Robust standard errors are estimated. ***, **, * denotes 1%, 5%, and 10% statistical significance.

	(1) Percentage Change Investment Homes	(2) Percentage Change Vacation Homes	(3) Percentage Change Stock Holdings
Constant	-0.08	-0.23	-0.37***
	(-0.52)	(-1.23)	(-4.20)
MSA Housing Return	1.14**	0.27	
	(2.19)	(0.29)	
MSA Local Stock Return			0.47
			(1.53)
Observations	40	33	41

Table 9: Explaining Housing Cycle of 2000-2010 Using Housing Supply Elasticity and log RATIO

We report estimates of the relationship of log RATIO and MSA Housing Elasticity on MSA home price appreciation, change in MSA price-to-rent ratio, and cumulative dollar returns investing in MSA housing market. The sample consists of MSAs with an average population (from 1996 to 2010) greater than 750,000. Panel A reports cross-sectional regression results using MSA home price appreciation as dependent variables. We use 4 quarter home price appreciation (t-4 to t) from quarterly MSA-level FHFA housing price. Columns (1) to (3) report the results using the cross-sectional average of variables from 2000 to 2006 and columns (4) to (6) report the results using the cross-sectional average of variables from 2007 to 2010. The independent variables are log RATIO, 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), MSA Housing Supply Elasticity, MSA unemployment rate, MSA income per capita, and log MSA Population. Due to the MSAs with RATIO equal to zero, we add 0.00001 to RATIO before we take log. Panel B reports cross-sectional regression results using % change in MSA price-to-rent ratio as dependent variables. We use the Zillow county-level home price index and the median county rent from the fair market rent, issued annually by the U.S. Department of Housing and Urban Development, to compute county-level Price-to-Rent ratio. Then we take the average of county-level Price-to-Rent ratio in an MSA to get MSA-level Price-to-Rent ratio. We use 4 quarter % change in MSA Price-to-Rent ratio (t-4 to t). Columns (1) to (3) report the results using the cross-sectional average of variables from 2000 to 2006 and columns (4) to (6) report the results using the cross-sectional average of variables from 2007 to 2010. The independent variables are same as in Panel A. Panel C reports cross-sectional regression results using the cumulative dollar returns investing in MSA housing market as dependent variables. Cumulative dollar returns are calculated using quarterly MSA-level FHFA housing index: from year 2000 1q to 2006 4q, and from year 2007 1q to 2010 4q. We use initial year values for the independent variables. That is, year 2000 values are used for columns (1) to (3) and year 2007 values are used for columns (4) to (6). The table reports point estimates with t-statistics in parentheses. ***, **, * denotes 1%, 5%, and 10% statistical significance. Economic Significance is the percentage change in the standard deviation of dependent variable implied by a one standard deviation increase in log RATIO.

Panel A: MSA Home Price Appreciation						
	(1)	(2)	(3)	(4)	(5)	(6)
Variables	From 2000 to 2006			From 2007 to 2010		
log RATIO	-0.0229***		-0.0137**	0.00783**		0.00275
	(-4.555)		(-2.479)	(2.054)		(0.693)
Economic Significance	-0.65		-0.39	0.25		0.09
MSA Housing Supply Elasticity		-0.0272***	-0.0199***		0.0187***	0.0174***
		(-4.675)	(-3.144)		(3.439)	(3.031)
Economic Significance		-0.59	-0.43		0.38	0.35
MSA Unemployment Rate	-0.00417	-0.000761	-0.00385	-0.0167***	-0.0155***	-0.0148***
	(-0.787)	(-0.152)	(-0.775)	(-4.914)	(-5.079)	(-4.573)
MSA Income Per Capita	2.50e-06**	-7.25e-07	8.60e-07	-9.82e-07	4.66e-07	2.19e-07
	(2.644)	(-0.871)	(0.842)	(-1.445)	(0.725)	(0.296)
log MSA Population	0.00401	-0.00661	-0.00231	-0.00520	-0.000203	-0.00122
	(0.609)	(-1.018)	(-0.357)	(-0.851)	(-0.0361)	(-0.208)
Constant	-0.0837	0.246**	0.110	0.213**	0.0297	0.0556
	(-0.873)	(2.517)	(1.020)	(2.453)	(0.340)	(0.583)
Observations	63	62	62	63	62	62
R-squared	0.267	0.280	0.351	0.462	0.510	0.514

Table 9 (continued)

Panel B: % Change in MSA Price-to-Rent Ratio						
	(1)	(2)	(3)	(4)	(5)	(6)
Variables	From 2000 to 2006			From 2007 to 2010		
log RATIO	-0.0191***		-0.0124**	0.00872*		0.00628
	(-4.114)		(-2.378)	(1.793)		(1.162)
Economic Significance	-0.61		-0.40	0.24		0.17
MSA Housing Supply Elasticity		-0.0227***	-0.0163***		0.0136*	0.0108
		(-4.156)	(-2.768)		(1.818)	(1.379)
Economic Significance		-0.55	-0.40		0.23	0.19
MSA Unemployment Rate	-0.00249	0.000415	-0.00247	-0.0173***	-0.0178***	-0.0162***
	(-0.512)	(0.0889)	(-0.531)	(-4.005)	(-4.242)	(-3.674)
MSA Income Per Capita	2.15e-06**	-5.49e-07	8.68e-07	-1.12e-06	1.24e-07	-4.39e-07
	(2.482)	(-0.712)	(0.913)	(-1.299)	(0.141)	(-0.438)
log MSA Population	0.00379	-0.00499	-0.00144	-0.00382	0.000867	-0.00132
	(0.626)	(-0.813)	(-0.236)	(-0.485)	(0.110)	(-0.163)
Constant	-0.105	0.169*	0.0531	0.168	0.0157	0.0730
	(-1.191)	(1.823)	(0.523)	(1.506)	(0.129)	(0.557)
Observations	62	61	61	62	61	61
R-squared	0.236	0.243	0.313	0.369	0.359	0.374

Panel C: Cumulative Dollar Returns Investing in MSA Housing Market						
	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Cumulative Dollar Return from 2000 to 2006			Cumulative Dollar Return from 2007 to 2010		
log RATIO	-0.239***		-0.135**	0.0643***		0.0352**
	(-4.210)		(-2.240)	(4.208)		(2.067)
Economic Significance	-0.59		-0.33	0.55		0.30
MSA Housing Supply Elasticity		-0.289***	-0.228***		0.100***	0.0780***
		(-5.136)	(-3.755)		(4.884)	(3.433)
Economic Significance		-0.55	-0.44		0.56	0.44
MSA Unemployment Rate	0.0722	0.1000**	0.0652	-0.0149	-0.0229	-0.0164
	(1.472)	(2.302)	(1.457)	(-0.864)	(-1.422)	(-1.029)
MSA Income Per Capita	2.70e-05***	-2.66e-06	9.53e-06	-5.32e-06**	3.18e-06	-2.22e-07
	(2.926)	(-0.323)	(0.988)	(-2.123)	(1.394)	(-0.0804)
log MSA Population	0.185***	0.0528	0.106	-0.0674***	-0.0357	-0.0464**
	(2.760)	(0.849)	(1.644)	(-2.929)	(-1.613)	(-2.095)
Constant	-2.424**	1.179	-0.158	2.198***	1.163***	1.516***
	(-2.525)	(1.293)	(-0.149)	(6.463)	(3.420)	(4.073)
Observations	63	62	62	63	62	62
R-squared	0.391	0.455	0.500	0.328	0.379	0.423

Figure 1: RATIO Variable across the U.S.

The figure shows the average RATIO variable across the metropolitan statistical areas in the U.S. with population greater than 750,000 within 1996 and 2010. RATIO is 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).

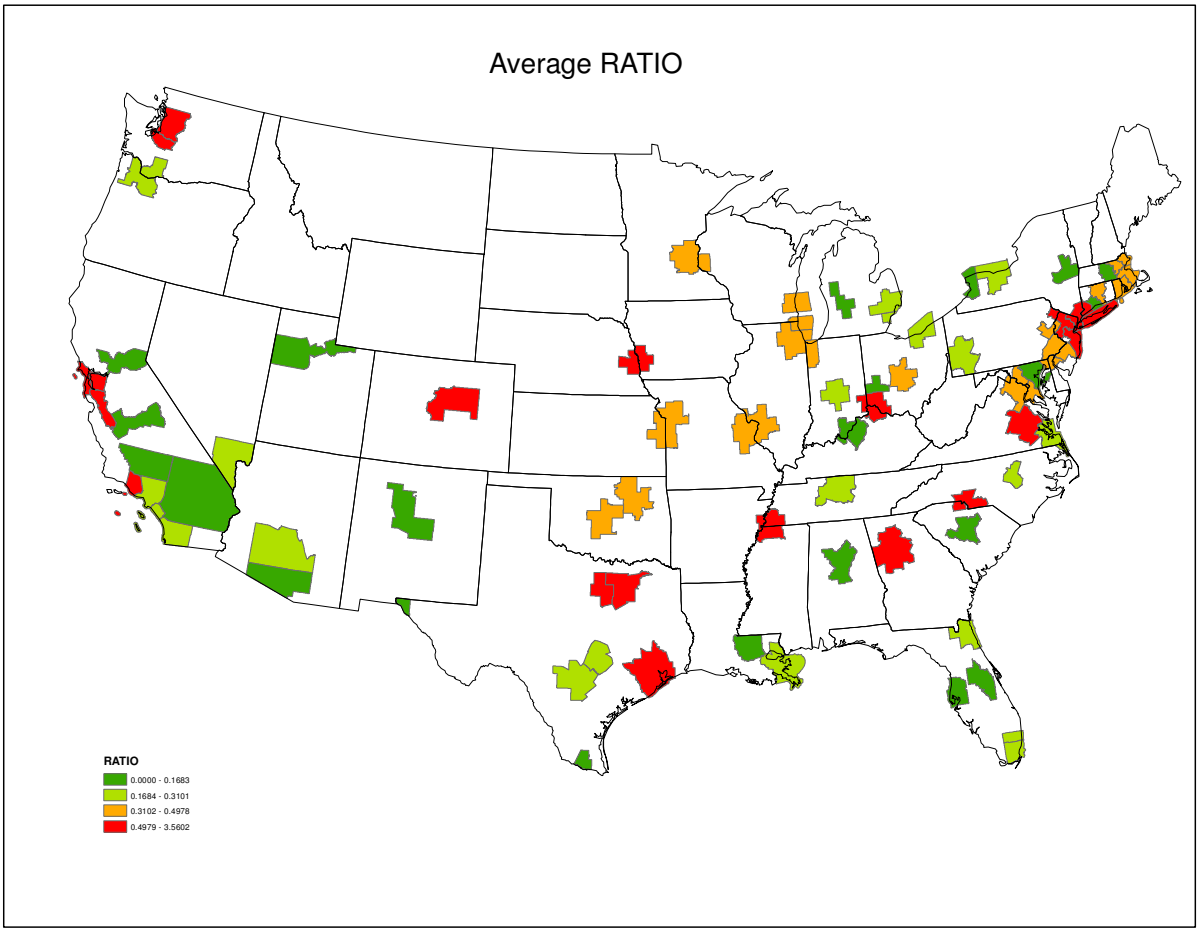

Figure 2: MSA/Year Residualized Scatter Plot

The figure shows scatter plots of the residualized MSA/YEAR log RATIO with the residualized MSA/YEAR investment real estate ownership (HAVE INVESTMENT REAL ESTATE) or the residualized MSA/YEAR % share of investment real estate in households' asset (% INVESTMENT REAL ESTATE IN TOTAL ASSETS). Panel A shows the scatter plot of residualized investment real estate ownership. Panel B shows the scatter plot of residualized % share of investment real estate in households' assets. Log RATIO is the log of 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.

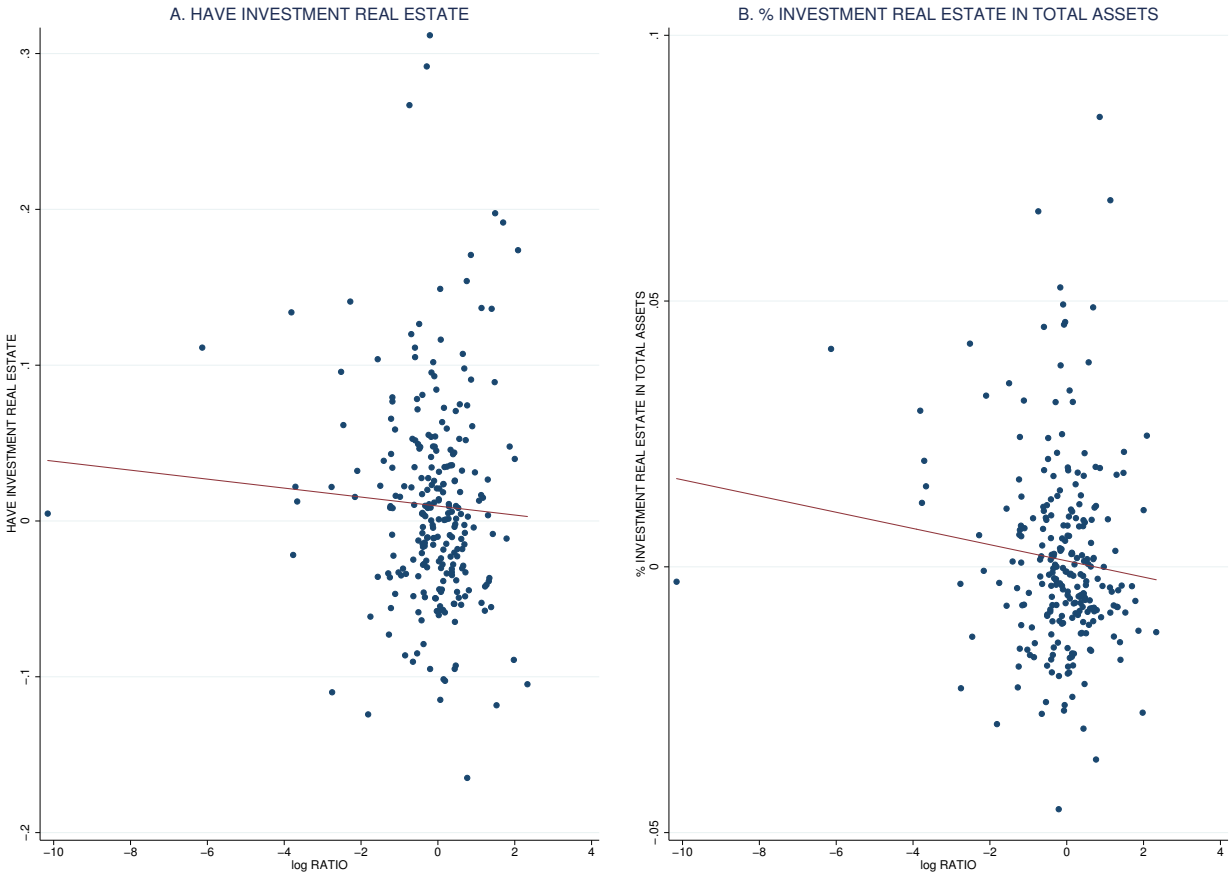

Figure 3: Housing Prices by Housing Supply Elasticity and by log RATIO

The figure shows the time series of median MSA home price appreciation, and median % change in MSA price-to-rent ratio by MSA Housing Supply Elasticity and by log RATIO. The sample consists of MSAs with an average population (from 1996 to 2010) greater than 750,000. Panel A shows the median MSA home price appreciation for 3 groups of elasticity and 2 groups of RATIO. We use 4 quarter home price appreciation (t-4 to t) from quarterly MSA-level FHFA housing price. The groups are independently sorted. For each elasticity group, blue solid line represents low RATIO (below median) group and red dashed line represents high RATIO (above median) group. Panel B plots the median % change in MSA Price-to-Rent ratio for 3 groups of elasticity and 2 groups of RATIO. We use the Zillow county-level home price index and the median county rent from the fair market rent, issued annually by the U.S. Department of Housing and Urban Development, to compute county-level Price-to-Rent ratio. Then we take the average of county-level Price-to-Rent ratio in an MSA to get MSA-level Price-to-Rent ratio. We use 4 quarter % change in MSA Price-to-Rent ratio (t-4 to t).

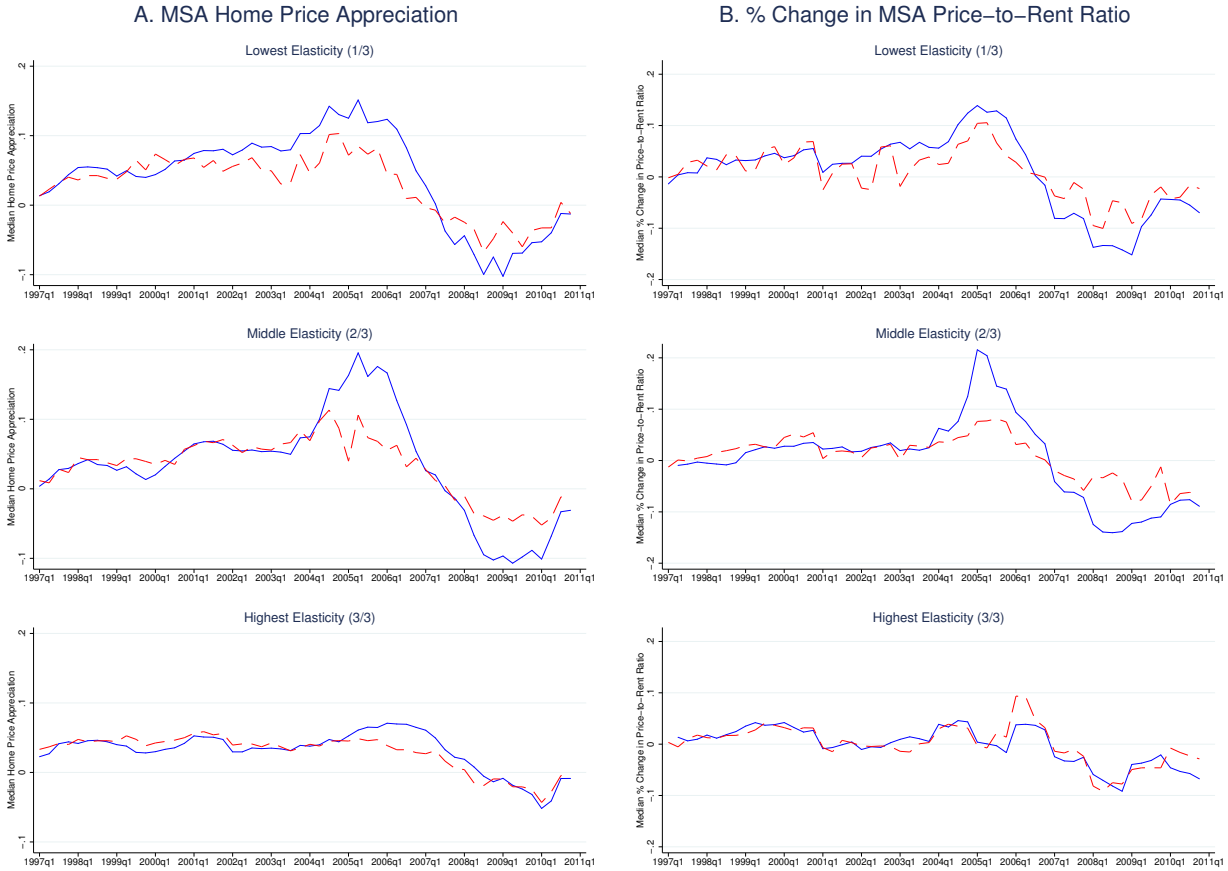

Figure 4: Cumulative Returns by Housing Supply Elasticity and by log RATIO

The figure shows median cumulative dollar returns investing in MSA housing market by MSA Housing Supply Elasticity and by log RATIO. We use quarterly MSA-level FHFA housing price. The sample consists of MSAs with an average population (from 1996 to 2010) greater than 750,000. Panel A shows the median of cumulative dollar return from 2000 1q to 2006 4q for 3 groups of elasticity and 2 groups of RATIO. The groups are independently sorted. For each elasticity group, blue solid line represents low RATIO (below median) group and red dashed line represents high RATIO (above median) group. Panel B shows the median of cumulative dollar return from 2007 1q to 2010 4q for 3 groups of elasticity and 2 groups of RATIO.