What Drove the 2003–2006 House Price Boom and Subsequent Collapse? Disentangling Competing Explanations*

John M. Griffin† Samuel Kruger‡ Gonzalo Maturana§

November 13, 2018

Ten years after the financial crisis, competing and often contradictory narratives have arisen around the central question of what can explain the massive rise and fall in house prices around the crisis. We provide a unified framework and use detailed cross-sectional data to examine four variants of the excess credit supply channel and three variants of the speculation channel that have been proposed in the literature. Although many proposed variables correlate well with house price patterns across regions, far fewer are consistently related to zip code level variation within regions both in the boom and the bust. The two variables that are selected by Bayesian model averaging as showing the strongest statistical and economic relation to house price changes, both in the boom and bust, are subprime lending and dubious origination practices. Surprisingly, none of the speculation measures including several possible extrapolative expectation proxies, explain zip code level house price growth variation within regions in both the boom and the bust. The effects of subprime lending and dubious origination are positively related to subsequent housing speculation and housing demand more generally. Inconsistent with lender expectations of future house price increases, credit supply is not correlated with house price growth in areas of elastic land supply; however, credit supply still predicts speculation growth, house transaction volume, and house price declines after the boom. Through both agency and non-agency loans, dubious lending practices seem to increase credit, and the effects of dubious lending are amplified in areas with likely income misreporting. Overall, our findings suggest that excess credit supply, particularly through subprime and dubious mortgage origination, stimulated housing demand and played a large role in the crisis.

JEL classification: G01, G21, R31

keywords: financial crisis, house price growth, credit supply, mortgage misreporting, housing demand, speculation

*We are grateful to Jonathan Cohn, Rohan Ganduri, Christoph Herpfer, Adam Kolasinski, Mattia Landoni, Richard Lowery, Gregor Matvos, Sheridan Titman, as well as seminar participants at Emory University, Texas A&M University, the University of Texas, the 2018 Lone Star Finance Conference, and the University of Chicago 10th Anniversary of the Financial Crisis Conference for helpful comments. We also thank Iman Dolatabadi and Lee Seltzer for excellent research assistance. We are grateful to the University of Texas McCombs Research Excellence Grant for research support and to Integra FEC for providing access to mortgage data. Griffin is an owner of Integra FEC and Integra REC which engage in financial consulting on a variety of issues including mortgage fraud. Integra REC is a party to RMBS litigation against banks. Supplementary results can be found in an Internet Appendix at the authors' websites.

†McCombs School of Business, University of Texas at Austin. Email: John.Griffin@utexas.edu.

‡McCombs School of Business, University of Texas at Austin. Email: Sam.Kruger@mccombs.utexas.edu.

§Goizueta Business School, Emory University. Email: Gonzalo.Maturana@emory.edu.

Ten years after the financial crisis, over 14,000 articles and 750 books have been written about the financial crisis.[1] Because there is widespread agreement that losses underlying the crisis initiated with house price losses, one of the most central questions regarding the financial crisis is what drove house prices. Surprisingly, there is little consensus on this question. There is considerable academic support for the view that house prices were driven by excess credit supply facilitated by originate-to-distribute securitization. However, there is also considerable support for the view that house price increases were driven by investor speculation. Additionally, there are considerable differences, debates, and contradicting narratives even within each of these main views. The purpose of this paper is to systematically examine the leading explanations in a uniform and objective empirical framework through which the evidence can be compared and contrasted.

The narratives converge around two main themes.[2] First, due to misaligned incentives, loan underwriting practices focused on originating loans to distribute and not to hold, resulting in a large expansion of credit (excess credit supply). The excess availability of mortgage credit shifted housing demand outward by enabling existing and new borrowers to spend more on housing. This caused house prices to rise and then crash when the loans could not be repaid and credit constraints tightened. Within this view, there are interesting variants. Was the excess credit focused on subprime borrowers (Mian and Sufi (2009)) or did securitization expand credit supply for borrowers more generally (Anenberg, Hizmo, Kung, and Molloy (2017); Mian and Sufi (2018))? Was the effect due to poor loan screening (Keys, Mukherjee, Seru, and Vig (2010); Purnanandam (2011)) for originate-to-distribute loans, or did cross-sectional differences in fraudulent origination practices also play a role in facilitating excess credit (Griffin and Maturana (2016a))?

The second major academic explanation is that the expansion in house prices was pri-

[1] On SSRN alone, 14,421 articles have been posted since 2007 with the phrase "financial crisis" in the title, abstract, or keywords. The book count is based on searching Amazon for books published since 2007 with the keyword "2008 financial crisis." A further manual review of book titles and Amazon book descriptions filtered out books not primarily concerned with the financial crisis, resulting in 759 published books.

[2] Mayer, Pence, and Sherlund (2009) and Levitin and Wachter (2012) survey explanations for the housing bubble.

marily driven by excessive investor optimism about future house price growth. Within this speculation view, there is also considerable variation across narratives. Was the expansion caused mainly by speculators in the form of non-occupant buyers (Haughwout, Lee, Tracy, and Van der Klaauw (2011); Gao, Sockin, and Xiong (2017)), out-of-town second home buyers (Chinco and Mayer (2016)), or land market speculators (Nathanson and Zwick (2018))? Theories also differ as to whether investors naively over-extrapolated past house price growth (Glaeser and Nathanson (2017); DeFusco, Nathanson, and Zwick (2017); Gennaioli and Shleifer (2018)) or experienced a more general shift in housing beliefs (Kaplan, Mitman, and Violante (2017)).

Most of the main narratives and sub-narratives have competing hypotheses and predictions. While the truth may be a combination of views, not all of the views can be equally true. The main task of this paper is to understand and evaluate the competing evidence for these explanations in the data. We primarily examine four credit supply and three speculation proxies that have been proposed in the literature. The four credit supply proxies are subprime share (Mian and Sufi (2009)), non-core-deposit liabilities (a measure of lender sensitivity to securitization growth) (Mian and Sufi (2018)), worse originator market share (a measure of the prevalence of lenders with dubious origination practices) (Griffin and Maturana (2016a)), and rate of private-label securitization. The three speculation and belief proxies are the percent of home purchases that are not owner-occupied (Gao, Sockin, and Xiong (2017)), out-of-town second-home purchaser share (Chinco and Mayer (2016)), and extrapolative beliefs based on past house price growth (Glaeser and Nathanson (2017); DeFusco, Nathanson, and Zwick (2017); Gennaioli and Shleifer (2018)).

We construct all measures as of 2002 and examine their relation to house price growth during the boom (2003–2006) and the bust (2007–2010). Our main analysis focuses on within-metropolitan statistical area (MSA) variation across zip codes, which controls for regional economic and housing shocks. Across MSA sorts provide support for most of the narratives, possibly explaining why there is still considerable disagreement in the literature.

We start with univariate analysis and then move to multivariate. From the univariate results, we first learn that most of the credit supply proxies are positively related to future house prices in the 2003–2006 boom and negatively related to the 2007–2010 bust. The relation between credit supply and house price growth during the boom is particularly strong in MSAs with low housing supply elasticity, consistent with Mian and Sufi (2009) findings for subprime zip codes. In areas with high housing supply elasticity, demand increases are offset by new housing construction (Glaeser, Gyourko, and Saiz (2008); Saiz (2010)). However, new housing construction can lead to excess housing supply. Consistent with this, credit supply during the boom is strongly related to price decreases during the 2007–2010 bust even in elastic areas.

Interestingly, none of the speculation proxies are related to prices in both the boom and bust. Non-owner occupancy is positively related to prices only in the boom, out-of-town purchasing is related to prices only in the bust, and 2002 and 2000–2002 house price growth are unrelated to prices in both the boom and the bust.

Multivariate specifications jointly assessing the proxies indicate that the credit supply proxies are jointly significant both on their own and after controlling for the speculation proxies. The most consistent credit supply variables are subprime share and worse originator share. Both of these variables are statistically and economically related to house prices in both the boom and bust after controlling for other credit supply proxies and speculation proxies. By contrast, the speculation proxies are jointly insignificant in the boom and are not significant in the bust after controlling for the credit supply proxies. Of the speculation proxies, non-owner occupancy performs the best in the boom (even after controlling for other proxies). However, none of the speculation proxies are significant in the bust after controlling for the credit supply proxies. Substantial robustness analysis, including weighted least squares, limiting credit supply proxy construction to purchase loans, and contemporaneously constructed measures, yields similar conclusions. Additionally, for extrapolative expectations we construct several additional measures of past house price growth based on

growth in a zip code over a three-year period (2000–2002), over a wider geographic areas spanning five miles around a zip code, and on a rolling annual basis each year from 2002 to 2005. None of these proxies consistently predicts subsequent within-MSA house price growth. Overall, it is striking that the speculation proxies receive such little support from within-MSA analysis.

We next turn to how credit supply may have fueled housing demand by examining how subprime share and worse originator share relate to subsequent housing speculation and transaction activity more generally. We find that these credit supply proxies are positively related to subsequent 2003–2006 non-occupant purchases as well as house purchase volume more generally. Combined with the regressions on boom and bust house price growth, this evidence indicates that subprime areas and areas with high worse originator share experienced credit supply expansion in 2003–2006 that increased housing demand and pushed prices up. Nonetheless, one might worry that subprime lenders and worse originators chose where to lend in anticipation of future house price growth or that their market share is correlated with an omitted variable related to future house price growth. To assess this possibility, we follow Mian and Sufi (2009) and focus on areas of high housing supply elasticity. Consistent with credit supply shocks increasing housing demand, subprime share and worse originator share predict non-owner occupancy and purchase transaction volume but not house prices during the boom in these elastic areas.

To understand the connection between worse originator share and credit supply expansion, we examine the role of misreporting in more detail. Originators with track records of high misreporting appear to have increased credit supply by relaxing lending standards. Zip codes with high worse originator share experienced increased mortgage originations despite declining income, and this effect is not limited to subprime areas. While worse originators are identified based on misreporting in mortgages that were privately securitized, the effect of worse originators is not limited to privately securitized mortgages. Zip codes with higher worse originator share also have higher delinquency rates on conforming Fannie Mae

mortgages. Income misreporting (Jiang, Nelson, and Vytlacil (2014); Ambrose, Conklin, and Yoshida (2016); Mian and Sufi (2017)) appears to be one of the ways in which lending standards were relaxed and could have facilitated an increase in demand as inflated income allowed borrowers to spend more on housing. Zip codes with high worse originator market share experienced high growth of rates of borrower-reported income compared to IRS income growth. In zip codes with high worse originator share, elevated differences between borrower-reported and IRS income growth are associated with larger 2003–2006 house price increases and larger 2007–2010 busts. This effect is not present in zip codes with low levels of originator misreporting.

We conclude the paper by looking at zip codes that had close to zero presence of worse originators and no evidence of income misreporting. These areas experienced modest house price growth in 2003–2006 followed by almost no decrease in 2007–2010. While this analysis falls short of being a clear counterfactual, it suggests that a world without misreporting may have looked considerably different.

Overall, our findings indicate that the subprime and worse originator share credit supply proxies strongly relate to house price changes during both the expansion and the crash. The results are consistent with credit supply growth fueling both speculation and demand for housing more generally. We discuss the relation between our findings and the literature in the next section. Our analysis casts significant doubt on the existing evidence for speculation and extrapolative expectations more generally. Although we use the leading proxies proposed over the last decade of research, it is possible that speculation measures are more poorly measured than proxies for credit supply. We mitigate this concern by considering multiple variants and robustness tests of speculation measures, but we cannot categorically rule out the speculation channel, and future research with stronger proxies could potentially alter conclusions. Our findings that excessive and fraudulent lending can have large distortive effects in real estate markets is consistent with academic research discussed in the next section and is also largely consistent with the crisis narrative of many financial practitioners

and journalists (Lewis (2010); Kolb (2010); McLean and Nocera (2010); Engel and McCoy (2011); Lowenstein (2011)).

1 Related literature

Since our paper fits within a recently developing literature that has examined potential explanations for house prices, we briefly describe the relevant features of that literature as well as the underpinnings that are relevant to our tests. When possible, variables from the previous literature are directly used in our tests.

1.1 Crisis narratives

The main narratives for house prices in the run-up to and aftermath of the financial crisis center around a shift in housing demand either driven by an increase in credit supply or optimistic investor views about housing.

The credit supply explanations build upon the "originate-to-distribute" view (Keys, Mukherjee, Seru, and Vig (2010); Purnanandam (2011); Keys, Seru, and Vig (2012); Rajan, Seru, and Vig (2015)). The main idea is that lenders prioritized loan origination volume for sale to securitization vehicles and the system malfunctioned by generating excess supply of low-quality loans that were increasingly securitized. When financing standards are relaxed, borrowers can take out more and larger loans, which can create an outward shift in the demand curve for housing, putting upward pressure on prices.[3] When the excess supply of credit is removed, the housing demand curve shifts back and house prices decrease. The underpinnings for this view are found in deteriorating underwriting standards (Demyanyk and Van Hemert (2011)).

A pure demand view says that investors may have had optimistic expectations for housing, which caused them to take out more credit. Shiller (2006, 2015) is an early and prominent

[3]House prices respond to a shift in the demand curve (Herring and Wachter (2000); Hubbard and Mayer (2009)). Pavlov and Wachter (2011) show theoretically that aggressive mortgage lending instruments magnify the real estate cycle.

advocate of this view. Arguments and models that speculation played a primary role in the housing crisis include Haughwout, Lee, Tracy, and Van der Klaauw (2011), Mayer (2011), Barlevy and Fisher (2011), Adelino, Schoar, and Severino (2016), Adelino, Schoar, and Severino (2018), DeFusco, Nathanson, and Zwick (2017), Kaplan, Mitman, and Violante (2017), Nathanson and Zwick (2018), and Gennaioli and Shleifer (2018).

A challenge to all narratives is finding empirical proxies that conceptually and analytically map to the individual channels. For our analysis, we build on the existing literature by implementing proxies proposed and tested in previous research within a common empirical framework.

1.2 Credit supply proxies

Mian and Sufi's (2009) paper is the most prominent test of the credit supply view. Mian and Sufi find that zip codes with large concentrations of residents with FICO scores below 660 as of 1996 experienced larger increases in mortgage origination and more house price growth from 2002 to 2005 than other zip codes despite having lower income growth. Nadauld and Sherlund (2013) find evidence that the increase in subprime lending is causally linked to the rise of securitization. In our empirical analysis, we construct a variant of the Mian and Sufi (2009) subprime zip code measure based on market share of subprime lenders that is also used by Adelino, Schoar, and Severino (2016) and Gao, Sockin, and Xiong (2017).

The subprime narrative has faced criticism because the increase in credit was not concentrated just in subprime zip codes. Adelino, Schoar, and Severino (2016) show that origination activity increased for middle income and wealthy borrowers and that these borrowers contributed substantially to defaults in the crisis. Foote, Loewenstein, and Willen (2016) and Albanesi, De Giorgi, and Nosal (2017) also show increases of debt for wealthy and prime borrowers. Similarly, Ferreira and Gyourko (2015) highlight the large incidence of prime borrower credit defaults.

What is often not clear in this discussion is that evidence against a subprime-only narra-

tive need not be inconsistent with the credit supply and originate-to-distribute views more generally. For example, Anenberg, Hizmo, Kung, and Molloy (2017) find that lenders increased maximum loan sizes from 2001 to 2006 for all borrowers, not just subprime borrowers, and that this increased house prices in MSAs most exposed to the credit expansion.[4] Mian and Sufi (2018) construct non-core-deposit liability ratio as a new credit supply measure that captures the extent to which banks rely on securitization for their lending as opposed to deposit financing.[5] Areas in which lenders have relatively high non-core-deposit liability ratios are more exposed to securitization expansion. We replicate this proxy in our empirical analysis as a general measure of exposure to credit supply expansion.

Another potential credit supply channel is that fraudulent practices may have increased credit supply in some geographic locations by enabling loans that would not have been otherwise approved. A large and growing academic literature shows that there was widespread mortgage fraud in the form of second-lien and owner occupancy misreporting (Piskorski, Seru, and Witkin (2015); Griffin and Maturana (2016b); Elul and Tilson (2016)), misreported income (Jiang, Nelson, and Vytlacil (2014); Ambrose, Conklin, and Yoshida (2016); Mian and Sufi (2017)), misreported assets (Garmaise (2015)), and inflated appraisals (Ben-David (2011); Kruger and Maturana (2018)). These practices led to over $137 billion in government legal settlements against banks. Piskorski, Seru, and Witkin (2015) and Griffin and Maturana (2016b) show that in addition to being widespread, there were significant differences in fraudulent mortgage origination practices across originators. Given that originator composition varies substantially with geography, geographic variation enables an empirical framework to assess whether originators who engaged in large amounts of misreporting distorted home prices. Griffin and Maturana (2016a) show that market share of originators with high levels of second-lien misreporting is associated with amplified house price growth

[4] Additional evidence that credit supply affected house prices includes Favara and Imbs (2015), Di Maggio and Kermani (2017), and Vojtech, Kay, and Driscoll (2016). Corbae and Quintin (2015) and Kermani (2012) propose models showing that an increase in credit due to lower standards contributed to the boom of housing prices and to the subsequent bust. In contrast, Glaeser, Gottlieb, and Gyourko (2012) argue that credit supply does not explain the house price boom.

[5] We thank Atif Mian and Amir Sufi for sharing this data.

and contraction. In our empirical analysis, we use and extend their measure.

Finally, following the intuition of Mian and Sufi (2018) and Nadauld and Sherlund (2013) that the increase in subprime lending was caused by securitization, we consider a zip code measure of the share of loans that are privately securitized.

1.3 Speculation and belief proxies

The main alternative to the credit supply view is that housing demand was driven not by excess credit supply, but by a direct shock to demand either in the form of a general shock to beliefs or through speculation by optimistic investors. Haughwout, Lee, Tracy, and Van der Klaauw (2011), Chinco and Mayer (2016), Adelino, Schoar, and Severino (2016), and Gao, Sockin, and Xiong (2017) all find that measures of speculation explain cross-sectional variation in house price growth.

Gao, Sockin, and Xiong (2017) show that areas with higher levels of speculation experienced greater house price increases and more housing construction during the 2004 to 2006 boom, followed by more severe outcomes during the subsequent economic downturn. Relatedly, DeFusco, Nathanson, and Zwick (2017) find that purchases by non-occupants were a large driver of increased transaction volume during this period. We follow Gao et al. and proxy for speculation with the percent of home purchases that are not owner-occupied in a zip code.[6]

Chinco and Mayer (2016) analyze twenty-one MSAs and find that MSAs with more out-of-town second-home purchasers experienced larger house price growth, indicating that out-of-town purchasers pushed up prices by behaving like misinformed speculators. We construct Chinco and Mayer's out-of-town purchaser measure at the zip code level using property deeds data using the same methodology but with more granular data over a wider sample of MSAs.

[6]Gao, Sockin, and Xiong (2017) also use state capital gains tax rates as an instrument for non-owner occupancy. We do not take this approach because it is not conducive to studying within-MSA house price differences.

A significant strand of the theoretical and empirical literature posits that real estate investors in part form extrapolative beliefs about future house price growth based on past house price growth (Glaeser and Nathanson (2017); DeFusco, Nathanson, and Zwick (2017); Gennaioli and Shleifer (2018)). As a result, house price growth exhibits positive serial correlation, and positive shocks to house prices can be followed by periods of excessive optimism. We proxy for this optimism in the cross section by using 2002 and 2000–2002 house price growth to predict subsequent house price growth.

1.4 Disentangling views

The credit supply and speculation views both specify channels that stimulate demand for housing and shift prices upward. Stylized facts about the crisis generally feature increases in both credit and speculative activity, and they could reinforce one another. For example, expansion of credit supply may have enabled speculation, and demand from speculators may have caused credit to expand.

Disentangling these views requires determining which is the predominate driver of house prices. To the extent that there is substantial geographical variation and sufficient power, proxies for credit supply and speculation should allow one to identify areas in which the speculation channel is active without excess credit supply and vice versa. Under the credit supply view, geographical variation credit supply expansion exposure should relate to future prices and housing demand, and this relation should hold after controlling for pre-existing exposure to speculation. Although speculative demand may be correlated with credit supply, the speculation view predicts that speculation proxies should relate to future house prices even after controlling for credit supply exposure. We have the benefit of starting with proxies for both credit supply and speculation that have been previously used in the academic literature and that had sufficient power to predict house prices and previous research. Nevertheless, like most empirical work, the precision of the empirical tests for both views is limited to the power of the empirical proxies. The regression framework we employ is subject to standard

concerns such as reverse causality, omitted variables, and measurement error. Nonetheless, empirical comparisons of proxies for each view are informative as to which narratives and sub-narratives most closely explain the rich cross-section of data regarding the rise and fall of house prices.

2 Data

We use detailed house price and loan data from Zillow, DataQuick, ABSNet, and the Home Mortgage Disclosure Act (HMDA), bank-level information from bank call reports, and demographic and economic data from the IRS and the U.S. Census Bureau.[7] Based on these data, we construct a detailed zip-code-level dataset of house prices and proxies related to exposure to credit supply and housing demand shocks during the 2003 to 2010 housing cycle. Because data coverage is better in metropolitan areas and most of our analysis controls for MSA fixed effects, we restrict the sample to zip codes within MSAs. As a starting point, Zillow has house price data for 11,444 zip codes within MSAs. After merging the other data sources, dropping zip codes with populations of less than 1,000, and requiring coverage for all supply and demand proxies, the final sample has 3,725 zip codes for variables calculated as of 2002 and 5,622 zip codes for explanatory variables calculated as of 2003–2006. Internet Appendices A and B describe the data and sample selection in more detail.

2.1 House prices

Our house price data are from Zillow, an online real estate data company. Zillow publishes house price indices based on median home values at different levels of aggregation, including metropolitan areas and zip codes.[8] We primarily analyze house price growth during two time

[7]DataQuick consists of residential property characteristics and transaction information based on county deed records. ABSNet consists of loan-level information on U.S. non-agency securitized mortgages based on servicer and trustee data. HMDA consists of mortgage application information which mortgage originators are required to report by law.

[8]A detailed description of their methodology is available at https://www.zillow.com/research/zhvi-methodology-6032/.

periods: 2003–2006 (defined as the change in the Zillow house price index from December 2002 to December 2006) and 2007–2010 (defined as the change in the Zillow house price index from December 2006 to December 2010). Table 1 reports summary statistics.[9] House prices increased by 43.5% during 2003 to 2006, which is equivalent to 9.1% per year. During 2007–2010, average zip code house price decreased by 18.9%.

[Insert Table 1 Here]

There is considerable variation in house price growth both across MSAs and within MSAs. During the 2003–2006 house price boom, the between-MSA house price growth standard deviation was 30.3% and the standard deviation across zip codes within MSAs was 13.9%. In 2007–2010, the between-MSA house price growth standard deviation was 16.8% with a within-MSA standard deviation of 7.8%. Highlighting this variance, Figure IA.1 of the Internet Appendix plots zip code house price growth in 2003–2006 and 2007–2010 with the well-known result that the house price cycle was significantly amplified in areas such as Arizona, California, Florida, and Nevada (the sand states) with more muted price dynamics in interior states. The plot also highlights the significant dispersion in house price growth within MSAs.

2.2 Proxies for credit supply and speculation

We construct zip-code level measures of each credit supply and speculation proxy based on detailed loan, transaction, bank, and house price data. Our primary focus is on proxies calculated as of 2002, but we also perform robustness with versions contemporaneous with the 2003–2006 boom period. As shown in Table 1, the proxies all exhibit significant variation both within and across MSAs. Additional details on how each measure is constructed are described in Internet Appendix C.

[9]The reported statistics for house price growth are for the larger sample of zip codes for which we have 2003–2006 data. Summary stats are similar in the sample 3,725 zip codes with 2002 proxies.

3 Relation between house prices and main measures

Before proceeding to detailed regression analysis, we first explore the general relation between house prices and the proxies for credit supply and housing demand. Because MSA house price growth is impacted by regional differences in economic growth, house supply elasticity, and other housing characteristics, we primarily focus on within-MSA analysis. Nevertheless, we also consider house price growth across MSAs.

Figure 1 plots house prices during 2003–2010 for geographic areas classified based on each proxy. The plots are indexed to December 2002 house prices, and areas are classified based on 2002 data. The plots on the left show average house prices for MSAs in the top and bottom quartile of each proxy. The plots on the right show average house prices for zip codes classified into within-MSA quartiles based on the same measures. The price fluctuations for high and low groups within MSA plots are generally smaller than differences across MSAs because price growth variation is smaller within MSAs.

[Insert Figure 1 Here]

Panel A of Figure 1 shows results for the credit supply proxies. Consistent with Mian and Sufi (2009), the within-MSA subprime plot shows that subprime zip codes experienced larger house price expansions in the boom and a larger contraction in the bust. However, subprime share is less related to house prices across MSAs. Consistent with Mian and Sufi (2018), the house price cycle was amplified in high non-core-deposit liability areas. The relation between non-core-deposit liabilities and house prices is particularly strong across MSAs but also shows up across zip codes within MSAs. Worse originator market share is also associated with an amplified house price cycle within and across MSAs, with a particularly strong relation across MSAs. The relation between private label securitization and house price growth is strong across MSAs during both the boom and the bust. Within MSAs, the relation is only present in the 2007 to 2010 bust.

Panel B of Figure 1 shows results for the speculation proxies. Consistent with Gao,

Sockin, and Xiong (2017), the relation between non-owner occupancy and house prices is strong across MSAs, but across zip codes within MSAs, there is a relation in boom but not the bust. The house price cycle was amplified in high out-of-town purchaser areas, with a particularly strong relation across MSAs. Finally, MSAs with large house price increases in 2002 continued to experience higher house price growth in 2003–2006, followed by severe contractions during 2007–2010. However, within-MSA house price growth differences in 2002 have no relation to subsequent house price growth.

Consistent with previous results in the literature, the plots in Figure **1** provide at least some support for all seven credit supply and speculation proxies. However, the proxies are potentially correlated, making it important to disentangle them from one another before interpreting the evidence as supporting any particular channel.[10] To differentiate between the proxies one needs a large number of observations, which is why we focus on zip-code level analysis with a cross section of 3,725 zip codes in 2002 and 5,622 zip codes in 2003–2006.

4 House price regressions

To disentangle the different explanations for the housing cycle, we regress house price growth during the boom (2003–2006) and bust (2007–2010) on proxies for each explanation in the cross section. Regressions are at the zip code level with MSA fixed effects and control for population, housing units, vacancy rates as of 2000, and income as of 2002.[11] The identifying assumption is that 2002 proxies capture exposure to subsequent unanticipated shocks during 2003–2006.[12] The main concern with this assumption is that the proxies are somehow correlated with anticipated differences in house price growth. While we cannot

[10]Table IA.**1** reports correlations between the proxies. Sixteen of the twenty-one pairwise correlations are positive and significant at the 5% level, and many are large.

[11]These control variables follow Griffin and Maturana (2016a) with the exception of excluding the contemporaneous control variable of income growth since all regression variables are prior to the 2003 to 2006 period. Population, housing units, and vacancy rates are from the 2000 decennial census. Average income is from IRS data.

[12]This is the same approach proposed by Mian and Sufi (2018) and is similar to the identification strategy used by Mian and Sufi (2009).

fully rule this out, in Figure IA.2 we plot house price appreciation from 2000 to 2003 for zip codes with low and high values of each proxy based on within-MSA quartiles. Unlike the price dispersion exhibited in Figure 1, price growth in 2000 to 2002 is unrelated to the 2002 credit supply proxies. For the speculation proxies, 2002 house price growth is related to 2000–2002 house price growth by construction. Non-owner occupancy and out-of-town purchases in 2002 also show some relation to 2000–2002 house prices. This suggests that the credit supply proxies capture exposure to unanticipated shocks as opposed to anticipation of expected price changes.

4.1 Individual proxies

The proxy variables are all standardized so that coefficients represent the effect associated with changing the proxy by one standard deviation. We first consider each proxy individually. Table 2 reports results for regressions of house price changes on individual proxies constructed using data as of 2002. In Panel A, the dependent variable is house price growth from 2003 to 2006, which was 46.3% on average. In Panel B, the dependent variable is house price growth from 2007 to 2010, which was -21.3% on average. t-statistics, reported in parentheses, are based on standard errors clustered by MSA.

[Insert Table 2 Here]

Columns (1) to (4) report results for the credit supply proxies. Consistent with Mian and Sufi (2009), 2002 subprime share is associated with higher house price growth during the expansion (panel A) and larger decreases during the subsequent contraction (panel B). Worse originator market share (column 3) and the rate of private securitization (column 4) are also associated with amplified house price growth and contraction. Non-core-deposit liabilities in 2002 have little relation to house price increases during 2003 to 2006.[13] However, non-core-

[13]This result differs from Mian and Sufi's (2018) finding that 2002 non-core-deposit liabilities predict subsequent house price increases mainly because including control variables reduces the effect of non-core-deposit liabilities. The Internet Appendix shows results consistent with Mian and Sufi (2018) with non-core-deposit liabilities relating to house price expansion in zip code regressions without controlling for average income (Table IA.2).

deposit liabilities are associated with larger house price declines during 2007 to 2010.

In the Internet Appendix (Tables IA.2 to IA.7), we consider alternative specifications without control variables, without MSA fixed effects, with an additional control variable for income growth, with weighted least squares (weighted by number of occupied housing units), and with alternative market share calculations for credit supply variables based on purchase loans only and weighted by loan size. Consistent with Figure 1, subprime share is unrelated to 2003–2006 house price expansion without MSA fixed effects. Non-core-deposit liabilities results are largely unaffected by weighted least squares and the alternative market share calculations, but the effect of non-core-deposit liabilities on 2003–2006 house price growth becomes significant when fixed effects or control variables are dropped. Private securitization is unrelated to house price expansion without control variables and when calculated based only on purchase loans. The relation between worse originator market share and house price changes holds in all specifications. Overall, the evidence strongly supports the worse originator share proxy and is reasonably supportive of subprime share for explaining within MSA, but not across MSA, variation. The private securitization and non-core-deposit liabilities proxies produce mixed evidence.

In addition to being highly statistically significant, the credit supply results in Table 2 are also economically meaningful. For example, a one standard deviation increase in worse originator market share is associated with an additional house price increase of 2.8 ppt during the boom and an additional decrease of 3.6 ppt during the bust. Relative to cross-sectional house price change standard deviations of 26.4% during the boom and 15.2% during the bust, these effects are substantial, especially considering that they come from a single noisy proxy for credit supply and are from within-MSA analysis, where there is less variation in house prices. Without MSA fixed effects (Table IA.3), a one standard deviation change in worse originator share has an even larger effect of 8.3 ppt during the boom and -6.3 ppt during the bust.

In columns (5) to (7) of Table **2**, we turn to proxies for speculation and house price expectations. Non-owner occupancy is associated with greater house price growth during 2003 to 2006, but there is no relation between non-owner occupancy and house price decreases during 2007 to 2010. The opposite is true for out-of-town purchasers. Incidence of out-of-town purchasers has no relation to house price increases during the 2003 to 2006 but is associated with larger declines during 2007 to 2010. This analysis differs from Chinco and Mayer (2016) and Gao, Sockin, and Xiong (2017) because it considers lagged measures of non-owner occupancy and out-of-town purchasers and is within MSAs.[14] We consider contemporaneous measures and results without MSA fixed below.

In contrast to the predictions of extrapolative expectations models, 2002 house price growth has no relation with subsequent within-MSA differences in house price growth. Because extrapolative expectations are a popular explanation for the house price bubble (Glaeser and Nathanson (2017); DeFusco, Nathanson, and Zwick (2017); Gennaioli and Shleifer (2018)), we consider this result in more detail in Table **3**. One possibility is that expectations are based on more than one year in house prices. Thus, we consider expectations based on three-year house price growth during 2000–2002 by regressing 2003–2006 house price growth on 2000–2002 house price growth in column (1). The estimated coefficient is negative, inconsistent with extrapolative expectations. Another possibility is that zip codes are too small. Perhaps expectations are formed over wider geographic areas. To assess this possibility, we recalculate 2002 and 2000–2002 house price growth based on all zip codes within a 5 mile radius of the zip code being analyzed.[15] As reported in columns (2) and (3), the broader house price growth measures also have no relation to subsequent house prices. A third possibility is that extrapolative expectations may operate at a higher frequency than one can identify solely with pre-2003 house price growth. To assess this possibility, we regress

[14] Chinco and Mayer (2016) analyze time-series and cross-sectional patterns during the boom period for twenty-one MSAs. Gao, Sockin, and Xiong (2017) analyze zip-code-level data but instrument for non-owner occupancy with state tax rates.

[15] Distances are from the center of one zip code to the center of other zip codes as identified by the US Census Bureau.

house price growth on past house price growth lagged by one year in columns (4) to (7). In these regressions, 2002 house price growth somewhat predicts 2003 house price growth, but 2003 house price growth has a negative relation with 2004 house price growth, and there is no relation between lagged house price growth and subsequent house price growth in 2005 or 2006.

[Insert Table 3 Here]

The bottom line from Table 3 is that we find no evidence for extrapolative expectations influencing house prices within MSAs. This contrasts with the predictive power of lagged house prices across MSAs described in Figure 1. What should one make of this difference? One possibility is that extrapolative expectations operate at the national or MSA level but not at more granular levels. However, there is nothing in the theory of extrapolative expectations that would predict this. Rather, the theory makes a general prediction that people form expectations about future house price growth based on past house price growth. This prediction applies equally well to hot and cold neighborhood, MSA, and national housing markets. Given that neighborhood house price expectations surely play a role in house buying decisions, extrapolative expectations should be just as relevant within MSAs as across MSAs. Thus, we interpret the lack of any relation between lagged house price growth and subsequent house price growth within MSAs as inconsistent with extrapolative expectations.

Figure 2 plots coefficients and 95% confidence intervals from Table 2 along with coefficient estimates from the same regressions without MSA fixed effects.[16] In general, coefficients are larger and confidence intervals are larger without MSA fixed effects. As discussed previously, subprime share does not predict boom house price increases without MSA fixed effects. The other credit supply proxies predict boom and bust house price changes without MSA fixed effects, as do non-owner occupancy and 2002 house price growth. Even without MSA fixed

[16]More detailed results for the regressions without MSA fixed effects are reported in Internet Appendix Table IA.3.

effects, out-of-town purchaser is insignificant in the boom and is only marginally significant in the bust.

[Insert Figure **2** Here]

4.1.1 Housing supply elasticity

Whether they come from credit supply increases or speculation, shocks to housing demand can be at least partially offset by new housing supply. As discussed by Glaeser, Gyourko, and Saiz (2008), in areas where new housing can be readily constructed, house prices should stay near construction costs even if housing demand increases. Mian and Sufi (2009) apply this logic to their zip code level subprime measure. Consistent with demand shocks being offset by new housing supply when construction is possible, subprime share has a larger effect in MSAs with low housing supply elasticity. We assess how housing supply elasticity interacts with all of the proxies for credit supply and speculation by separately estimating the effect of each proxy on house prices in high and low housing supply elasticity MSAs. Housing supply elasticity is from Saiz (2010), and the sample is split based on the median elasticity across zip codes.

Panel A of Figure **3** plots the results from regressions of 2003–2006 house price growth on individual proxies separately for low elasticity MSAs and high elasticity MSAs. All of the credit supply proxies predict 2003–2006 house price growth in low housing supply elasticity MSAs. By contrast, credit supply has no relation with 2003–2006 house price growth in high housing supply elasticity MSAs, consistent with new construction offsetting housing demand in these areas. Though the differences are not as large, the pattern is similar for non-owner occupancy.

[Insert Figure **3** Here]

While housing demand shocks in high housing supply elasticity areas should lead to little to no house price growth, they could still be associated with amplified house price declines

during the bust period if new housing construction creates excess housing capacity. This is exactly what we find in regressions of 2007–2010 on 2002 credit supply proxies plotted in Panel B. The credit supply proxies are associated with larger house price decreases during the bust in both high and low housing supply elasticity areas. In the Internet Appendix (Table IA.8), we also split out sand states (Arizona, California, Florida, and Nevada) from the rest of the country and find that credit supply's effect on house prices is concentrated in the sand states during the boom but is widespread during the bust.

4.1.2 Contemporaneous proxies

The 2002 credit supply and speculation demand proxies have the advantage of being plausibly exogenous to shocks in 2003 to 2006. However, they are likely noisier than contemporaneous measures as the housing market changed considerably from 2002 to 2006. Thus, we complement our 2002 proxy analysis with analysis of credit supply and speculation proxies calculated with 2003–2006 data.[17] Coefficients and confidence intervals from regressions of house price changes on individual contemporaneous proxies with and without MSA fixed effects are plotted in Figure 4. All regressions include the same control variables used in the lagged regressions. More detailed regression results are reported in Internet Appendix Tables IA.9 and IA.10.

[Insert Figure 4 Here]

Using contemporaneous data, all four credit supply measures are strongly related to house prices in both the boom and the bust with and without MSA fixed effects. For the speculation measures, non-owner occupancy is significantly related to price changes in both the boom and bust. Somewhat surprisingly, out-of-town purchasers is negatively related to price increases in the boom with MSA fixed effects and is insignificant in all other specifications. This contrasts with the positive relation that Chinco and Mayer (2016) find between out-of-town

[17] Because extrapolative expectations are inherently backward-looking, we do not have a contemporaneous version to the 2002 house price growth proxy.

purchases and house price growth during the boom period and provides additional evidence to indicate that the Chinco and Mayer effect is not robust.

The economic magnitude of the coefficients in the contemporaneous specifications are considerably larger than for the 2002 proxies, consistent with the lagged proxies being somewhat noisy. For example, a one standard deviation increase in worse originator share corresponds to a 5.9 ppt increase in house price growth during 2003–2006 and a 5.2 ppt decrease in house prices in 2007–2010, compared to estimates of 2.8 and -3.6 ppt found in Table **2**.

Overall, a rather surprising bottom line from the analysis of individual proxies is that within-MSA analysis of 2002 proxies lends support for only some of the proxies proposed in the literature. Subprime share, worse originator share, and (to a lesser extent) private securitization rate predict subsequent booms and busts in house prices as expected. Non-owner occupancy predicts increases during the boom but not decreases during the bust. Non-core-deposit liabilities and out-of-town purchaser predict the bust but not the boom. And, inconsistent with extrapolative expectations, 2002 and 2000-2002 house price growth have no relation to subsequent house price changes in the boom or the bust. Contemporaneous proxies for non-core-deposit liabilities and non-owner occupancy perform better than their lagged counterparts, with expected amplification effects during both the boom and the bust. However, contemporaneous out-of-town purchaser has no relation to house price declines during the bust and has the wrong sign during the boom.

4.2 Multivariate analysis

To disentangle the proxies, we turn to multivariate analysis to see how each proxy relates to house price growth after controlling for the other proxies. Table **4** reports results for the 2002 proxies. As before, panel A reports results for 2003–2006 house price growth, and panel B reports results for 2007–2010 house price growth. Regressions control for MSA fixed effects and the same control variables used previously. Column (1) reports results for the credit supply proxies. F-tests for joint significance indicates that the credit supply proxies

have a significant effect on house prices during both the boom and the bust. Subprime share and worse originator share predict house price increases in 2003–2006 and house price declines in 2007–2010, even after controlling for the other credit supply proxies. Non-core-deposit liabilities are significantly related to house price decreases during the bust but are again unrelated to house price growth during the boom. Private securitization is unrelated to subsequent house price changes in either period after controlling for the other proxies. Overall, subprime and worse originator share have the most consistent relation to house price growth among the credit supply proxies. In addition to being statistically significant, the coefficients on these variables are economically meaningful. A one standard deviation increase in worse originator share is associated with a 1.9 ppt increase in house prices during 2003 to 2006 and a 2.0 ppt decrease in house prices during 2007 to 2010.

[Insert Table 4 Here]

Column (2) of Table 4 reports multivariate results for the speculation proxies. Consistent with the analysis of individual proxies in Table 2, housing demand proxies have inconsistent relations with house price growth. F-tests indicate that the speculation proxies are jointly significant during the bust but are insignificant during the boom. Non-owner occupancy is associated with house price increases in 2003–2006 but has no relation to house price declines in 2007–2010. Out-of-town purchaser is related to house price declines in 2007–2010 but has no relation to house price increases in 2003–2006. House price growth in 2002 has no relation to subsequent house price changes in either period.

Column (3) of Table 4 jointly considers all seven proxies for credit supply and speculation. F-tests indicate that the credit supply proxies are significant in both periods even after controlling for all of the speculation proxies. Once again, subprime share and worse originator share are the most consistent credit supply proxies, though significance in the boom period decreases somewhat. In short, controlling for speculation has no effect on the credit supply evidence. In contrast, the speculation proxies are jointly insignificant in both the boom and the bust after controlling for the credit supply proxies. Individually, non-owner occupancy is

significant in the boom. All other coefficients are insignificant. Column (4) employs variable selection with the Bayesian information criteria to select optimal predictive models. The selected variables in the boom period are subprime share, worse originator share, and non-owner occupancy. In the bust, subprime share, non-core-deposit liabilities, worse originator share, and 2002 house price growth are selected, with an insignificant coefficient for 2002 house price growth.

We also consider the Bayesian model averaging approach to model selection, which estimates posterior distributions for both the probability that a variable is included in the model and the variable's coefficient conditional on inclusion. As discussed by Fernandez, Ley, and Steel (2001) and Ley and Steel (2009), Bayesian model averaging outperforms individual model selection and yields inferences on inclusion probabilities and posterior distributions that are lacking in standard variable selection.[18] We follow the modeling and prior distribution assumptions recommended by Ley and Steel (2009). Posterior inclusion probabilities are reported in column (5) of Table 4, and posterior coefficient distributions are plotted in Figure 5. Consistent with the rest of Table 4, Bayesian model averaging results in high posterior inclusion probabilities for subprime share, worse originator share, and non-owner occupancy with posterior coefficient distributions that are almost entirely above zero for 2003–2006 house price growth. For 2007–2010 house price growth, the most important coefficients are subprime share, non-core-deposit liabilities, worse originator share, and 2002 house price growth.

[Insert Figure 5 Here]

In the Internet Appendix (Table IA.11), we repeat the multivariate analysis of Table 4 with contemporaneous proxies. The credit supply proxies are again jointly significant in both periods even after controlling for the speculation proxies. The contemporaneous speculation proxies are jointly significant in the boom but are insignificant in the bust after

[18] See also Varian (2014) for a less technical discussion of Bayesian approaches to model selection.

controlling for the credit supply proxies. Similar to the univariate results, the magnitudes of the contemporaneous non-owner occupancy coefficients are about half of those of the credit supply proxies.

Overall, much can be learned from putting all of the proxies on equal footing. For the most part, the speculation proxies line up well with house price data when aggregated to the MSA level but are generally not able to explain zip code level variation within MSAs. In particular, none of the speculation proxies measured as of 2002 are consistently related to house prices in both the boom and the bust. In particular, 2002 house price growth has no relation to subsequent within-MSA house price growth during the boom or the bust. Similarly, the contemporaneous Chinco and Mayer out-of-town purchaser variable is unrelated to within-MSA house price variation both in the boom or bust, even before comparing it to other measures.

With respect to credit supply, all of the measures are strongly statistically and economically related to the bust in house prices in univariate specifications. Though correlated, subprime share and worse originator share as of 2002 are the most consistently related to within-MSA house price growth variation in both the boom and bust. The statistical and economic significance of these credit supply proxies are considerably larger in the bust.

4.3 Credit contraction

While most of our analysis follows the previous literature and focuses on how proxies for credit supply and speculation exposure relate to the full housing cycle, credit contraction may have also had a direct impact on house prices during the 2007–2010 bust. Panel A of Figure **6** plots lending volume by year based on HMDA data. Overall lending volume averaged $2.66 trillion per year from 2002 to 2006 and then contracted by 45.8% from 2006 to 2008. This contraction was particularly pronounced for non-bank mortgage companies, which experienced a 66.0% contraction, including many high profile collapses such as New Century's bankruptcy in April of 2007.

[Insert Figure 6 Here]

Because mortgage company market share varies across geographical areas, the collapse of mortgage company lending affected some areas more than others. To assess the impact of this contraction, we sort zip codes based on mortgage company share and match zip codes in the highest quartile of mortgage company market share to zip codes in the lowest quartile of mortgage company market share in the same MSA that experienced similar house price growth in 2003–2006.[19] By matching on 2003–2006 house price growth, we control for housing market conditions during the boom and focus on the impact of credit contractions on house price decreases during the bust. Panel B plots monthly house prices for high and low mortgage company share zip codes. By construction, house price growth is the same for both samples during the boom. However, as mortgage company lending contracts during the bust period, zip codes with high mortgage company share experienced substantially larger house price declines. In 2007–2010 zip codes with high mortgage company market share experienced house price declines of 27.0%, compared to 17.7% for zip codes with low mortgage company market share.

To assess statistical significance and control for other zip code characteristics, we also regress 2007–2010 house price changes on 2006 mortgage company market share across all zip codes controlling for 2003–2006 house price growth, MSA fixed effects, and the same zip code characteristics used in previous regressions. To see how the impact of 2006 mortgage company market share on house prices evolves over time, we estimate separate regressions of cumulative house price changes from the end of 2006 until subsequent quarters. Panel C of Figure 6 plots the resulting coefficients along with 95% confidence intervals based on standard errors clustered by MSA. A one standard deviation increase in 2006 mortgage company share is associated with a significantly larger house price declines. The difference in house prices materializes quickly and is persistent throughout the bust period, reaching -3.8

[19]Specifically, we require matched zip codes to have a difference of at least 15% in mortgage company share and a difference of no more than 1% in 2003–2006 house price growth. This results in a matched sample of 331 zip codes.

ppt by the end of 2010. Regression results for the full 2007–2010 time period are reported in Table IA.12.

5 Relation between credit supply and housing demand

As discussed by Mian and Sufi (2009, 2018), greater availability of mortgage credit allows households to spend more on housing and enables additional households to enter the market. This increases housing demand, which puts upward pressure on house prices. Because credit supply impacts house prices through housing demand, we next turn to analyzing how the subprime share and worse originator share credit supply proxies are related to housing demand proxies.

We begin investigating whether credit supply expansions increased speculative demand, proxied by the share of house purchases in 2003–2006 that are non-owner occupied. Results are presented in Table 5. In columns (1) and (2), specifications are the same as in the previous tables, except that the dependent variable is 2003–2006 non-owner occupied house purchase share as opposed to house price growth. In addition, because non-owner occupancy is likely be highly persistent, we include its 2002 level as a control variable. The regressions show that 2002 subprime share and worse originator share positively predict 2003–2006 non-owner occupancy. A one standard deviation increase in 2002 subprime (worse originator) share is associated with a 0.5 (0.4) ppt increase in 2003–2006 non-owner occupancy relative to the mean non-owner occupancy rate of 13.4% and cross zip code standard deviation of 10.5%.

[Insert Table 5 Here]

As a more general measure of market activity, we next assess the relation between the credit supply proxies and house purchase volume, which we compute as the annual number of purchase transactions per zip code in DataQuick property deeds data divided by the existing number of residential properties in the zip code. Mian and Sufi (2018) similarly analyze

transaction activity as a measure of demand and find that it is positively related to non-core-deposit liabilities.[20] Do credit supply proxies predict subsequent housing transaction activity? In columns (3) and (4) of Table **5**, we regress 2003–2006 house purchase volume on 2002 credit supply proxies. The regressions control for 2002 volume to ensure that the results are not from pre-existing correlations between volume and the credit supply proxies. A one standard deviation increase in 2002 subprime share is associated with 0.4 ppt increase in house purchase volume in 2003–2006 relative to a mean volume level of 8.4% of the housing stock and cross zip code standard deviation of 3.2%. Similarly, a one standard deviation increase in 2002 worse originator market share is associated with a 0.3 ppt increase in 2003–2006 volume.

Following DeFusco, Nathanson, and Zwick (2017) we decompose volume into short term and long term volume based on the length of time since a property was last sold. Specifically, we define transactions as short term if the property being sold was purchased within the last year, and we classify all other transactions as long term.[21] As discussed by Adelino, Schoar, and Severino (2016), DeFusco et al., and Mian and Sufi (2018) short-term volume likely reflects speculation and was the main driver of the overall increase in purchase volume during the housing boom. If credit supply increased speculation, we would expect credit supply proxies to predict short-term volume. This is exactly what we find in columns (5) and (6) of Table **5**. Subprime share and worse originator share in 2002 predict subsequent short-term volume with coefficients that are similar to overall volume and larger than long-term volume despite the much lower mean level of short term volume compared to long term volume (1.3% of the housing stock compared to 7.1% of the housing stock).

The evidence in Table **5** indicates that credit supply facilitated non-owner occupancy pur-

[20] While the literature typically interprets house purchase volume as a demand variable, it could also reflect supply changes. To assess whether volume changes are more consistent with demand or supply changes, we regress house price changes on house purchase volume in Internet Appendix Table IA.**13**. Consistent with volume reflecting demand, 2002 and 2003–2006 house purchase volume positively relate to price increases in the boom and negatively relate to prices in the bust.

[21] DeFusco, Nathanson, and Zwick (2017) consider different lengths of time and primarily focus on a cutoff of three years for defining a transaction as short term. Because our transaction data starts in 2002, we are not able to calculate longer windows, nor are we able to decompose 2002 transactions.

chases as well as demand for house purchase transactions more generally. Combined with Section 4's house price growth results, the evidence strongly supports the credit supply channel. Subprime areas and areas with high worse originator share experienced credit supply expansion in 2003–2006 that increased housing demand and pushed prices up. Nonetheless, one might worry that subprime lenders and worse originators chose where to lend in anticipation of future house price growth or that their market share is correlated with an omitted variable related to future house price growth. To assess this possibility, we examine the relation between credit supply and housing demand in a setting in which house prices did not increase. As discussed in Section 4.1.1, high housing supply elasticity areas provide such a setting. Areas with high supply elasticity respond to demand shocks with additional housing construction, which keeps house prices increases minimal.

How did non-owner occupancy purchasing and house purchase turnover respond to credit supply in high elasticity areas? Table **6** replicate Table **5**'s regressions in a restricted sample of only MSAs with high housing supply elasticity.[22] Even though credit supply did not increase house prices in these areas, 2002 subprime share and worse originator share are both positively related to subsequent non-owner occupancy, overall purchase volume, and short-term volume.[23]

[Insert Table **6** Here]

6 Misreporting and house prices

Given the important role that worse originator share played in the house price expansion and collapse, we next turn to investigating misreporting in more detail.

[22] As discussed in Section 4.1.1, housing supply elasticity is at the MSA level from Saiz (2010), and the sample is split at the median across all zip codes.

[23] Table IA.**14** reports results for low housing supply elasticity areas. The credit supply proxies positively predict 2003–2006 volume and short-term volume in low housing supply elasticity MSAs. However, there is no relation between the credit supply proxies and subsequent non-owner occupancy in these areas.

6.1 Mechanism

Why did house prices increase more in areas in which originators with high incidences of misreporting had larger market shares? Griffin and Maturana (2016a) hypothesize that misreporting originators increased the credit supply by underwriting low quality loans, causing an expansion of credit in zip codes with high worse originator presence.

We have already seen that zip codes with high market share for originators with high rates of second-lien misreporting experienced larger house price increases and elevated housing demand. To more directly assess whether worse originators expanded mortgage credit supply, we regress growth in mortgage credit on worse originator share. Table 7 reports the results of regressions with MSA fixed effects and the same control variables as in previous regressions. As reported in column (1), a one standard deviation increase in worse originator share is associated with an 8.3 ppt increase in mortgage growth relative to a 14.5% average increase in annual mortgage volume from 2002 to 2003–2006. Column (2) reports results from a regression of IRS income growth on worse originator share. The coefficient associated with worse originator share is negative. Zip codes with high worse originator share in 2002 experienced less income growth in 2003 to 2006 than other zip codes. This is similar to the credit supply evidence Mian and Sufi (2009) find for subprime zip codes and is the opposite of what one would expect to find if the worse originators were somehow anticipating a demand shock.

[Insert Table 7 Here]

One question is whether this effect is driven by subprime zip codes. Columns (3) and (4) repeat the same regressions in a sample restricted to zip codes with subprime lender shares that are below the median. Worse originator market share is strongly associated with mortgage growth even in these low subprime zip codes. Therefore, worse originator credit expansion was not restricted to subprime areas. Rather, the effect was similarly large in non-subprime areas. This potentially helps reconcile some important findings in the literature.

Adelino, Schoar, and Severino (2016), Foote, Loewenstein, and Willen (2016), and Albanesi, De Giorgi, and Nosal (2017) find that the growth in credit was not limited to subprime areas and infer from this that the growth in credit is not due to the supply channel. While it is true that these findings are inconsistent with Mian and Sufi's (2009) subprime channel being the sole driver of credit growth, subprime lending is just one proxy for credit expansion. The above findings indicate that poor origination practices were strongly related to credit growth in non-subprime areas as well.

Why did zip codes with high worse originator share experience credit growth? Did zip codes with high worse originator share experience relaxed lending standards? If so, were relaxed standards limited to privately securitized loans or did they also affect agency loans? Griffin and Maturana (2016b) find that non-agency loans originated by worse originators default at higher rates. To assess whether worse originator market share is related to delinquencies for higher-quality conforming loans, we analyze the relation between delinquency and worse originator market share in Fannie Mae data for loans originated between 2003 and 2006. For consistency and due to data availability, we calculate worse originator share based on privately securitized mortgages in the ABSNet data as before. Two limitations of this analysis are that Fannie Mae data only includes three digit zip codes and only reports originator names for originators with market share over 1%. Thus, we use worse originator share at the three-digit zip code level and group smaller lenders together for purposes of calculating lender fixed effects. Results are reported in Table **8**. Consistent with privately securitized mortgages, worse originator share is associated with higher delinquency rates with a coefficient of 0.7 ppt, relative to a mean delinquency rate of 7.7% (column (1)). This coefficient remains unchanged when adding lender fixed effects to the specification (column (2)). The results indicate that areas with higher worse originator market share experienced more delinquencies even after controlling for MSA×quarter of origination fixed effects and lender fixed effects for large lenders.

[Insert Table **8** Here]

In column (3), we add the interaction between worse originator market share and an indicator for loans originated by unidentified originators with less than 1% market share. Consistent with worse originators representing a larger share of the unidentified originators in zip codes with high worse originator market share, the interaction coefficient is positive and significant. Overall, the results suggest that the effects of dubious origination practices were present in both agency and non-agency loans.

6.2 Income misreporting

One of the most striking pieces of evidence of subprime credit expansion provided by Mian and Sufi (2009) is that mortgage expansion and price growth in subprime zip codes were accompanied by decreasing IRS income as opposed to increasing income as housing demand growth would predict. Adelino, Schoar, and Severino (2016) question this evidence by noting that HMDA income grew in subprime areas, consistent with housing demand growth. The conflicting evidence from HMDA and IRS income growth sparked a debate over whether differences between IRS and HMDA income growth should be interpreted as resulting from income misreporting in mortgage applications underlying the HMDA data (Mian and Sufi (2017)) or as evidence of changes in the relative incomes of homeowners and renters within a zip code (Adelino, Schoar, and Severino (2016)).

We calculate income growth during 2003 to 2006 as the annual growth rate in average income from 2002 to 2006. Because HMDA income comes from reported income on mortgage applications and IRS income is the average income of all households in a zip code, the difference in their growth rates conceptually captures both shifts in the income of homeowners relative to other households and also shifts due to income misreporting. It is an empirical question as to which of these effects dominates.

Areas with a large presence of worse originators provide a laboratory for investigating the income misreporting hypothesis. If HMDA-IRS income growth differences are due to a shift in composition of who owns houses in a zip code, there is no reason for income growth

differences to be related to worse originator market share. By contrast, to the extent that worse originator market share relates to unscrupulous lending practices, areas with high worse originator market share may also have high income misreporting. We first test this hypothesis by regressing income growth difference on worse originator market share.[24] As reported in column (1) of Table 9, zip codes with higher worse originator market share in 2003–2006 experienced larger differences between HMDA and IRS data on income growth from 2003 to 2006.

[Insert Table 9 Here]

The relation between worse originator market share and HMDA-IRS income growth differences suggests that income growth differences are at least in part related to misreporting. Nonetheless, there could still be zip codes in which the difference is due to a shift in the composition of home buyers. HMDA-IRS income growth differences may reflect misreporting in areas with high worse originator share but also could be due to demographic changes in areas with low misreporting.

To more thoroughly examine this issue, we jointly regress house price changes on worse originator share and HMDA-IRS income growth difference. Columns (2) to (7) of Table 9 report the results. Income growth differences are associated with increased house price growth during the boom (column (2)) and larger decreases during the bust (column (5)). Consistent with income growth differences capturing misreporting in areas with unscrupulous lending practices, income growth differences have less impact on 2003–2006 house prices after controlling for worse originator share in column (3). Additionally, the interaction term between worse originator share and income growth difference is positive in column (4). After controlling for worse originator share, income growth difference no longer has a significant relation to house price decreases during the bust in columns (6) and (7).

[24]Because we are not focused on causal statements but on contemporaneous relations, to put all variables on equal footing, we measure all variables from 2003 to 2006.

We graphically examine this relation in Figure 7. Zip codes with high income growth differences between HMDA and IRS data in 2003–2006 experienced more house price growth in 2003–2006 and larger house price decreases in 2007–2010, but this relation only holds in zip codes that had high worse originator share. In zip codes within the lowest three quintiles of worse originator share, there is no relation between income growth difference and house price growth during either the boom or bust period.[25]

[Insert Figure 7 Here]

Changes in HMDA-IRS income growth differences conceptually capture shifts in the income distribution and also shifts due to misreporting. In sum, we see evidence for both patterns in the data. However, we see little relation between income growth differences and house prices in the zip codes where misreporting is less likely. In zip codes with a higher composition of questionable lending practices, the income growth difference measure is more likely to be related to misreporting, and we see a strong relation between income growth differences and house prices. These results reinforce the point that all of our proxies are noisy and cannot perfectly measure issues such as misreporting. Nevertheless, when both proxies are present, the house price swings are largest. In particular, in zip codes with both proxies in the top quintile, house prices increased by 57% in 2003–2006 and decreased by 28% in 2007–2010. By contrast, in the zip codes in the bottom quintile of both proxies, the swings are 38% and -15%. This 50% relative increase in house price growth during the boom and doubling of house price contraction during the bust suggests that excess supply from misreporting played a sizeable role in house price swings. Moreover, this analysis is entirely within MSAs, which likely understates the role of misreporting. In the next section, we consider zip codes with no evidence of misreporting across all MSAs.

[25]The Internet Appendix (Figure IA.3) shows the time-series plots of the low and high worse originator share with income growth difference that are behind the result.

6.3 House price growth and contraction without misreporting

One of the hardest questions to answer is what house prices would have been in the absence of fraud and misreporting. This is challenging because it is difficult to construct counterfactuals in a world where an activity is widespread and correlated with other outcomes.

While it falls short of being a clear counterfactual, it is interesting to look at house prices in areas that were relatively unaffected by misreporting. Figure 8 plots house prices over time in areas with different levels of misreporting exposure. The blue circles plot house prices in zip codes with worse originator share over 5% in 2003–2006 and HMDA-IRS income growth difference exceeding 2.5% (both of which correspond to approximately the 75th percentile), the dotted line plots house prices in zip codes with worse originator share over 5%, the dashed line plots house prices in zip codes with worse originator share of less than 1%, and the hollow triangles plot house prices in zip codes that have both worse originator market share of less than 1% and did not experience any higher income growth in HMDA data than in IRS data. Whereas the zip codes with high worse originator share and high income growth differences experienced house price growth of 69% during 2003–2006, prices rose only 26% in zip codes with low worse originator share and only 23% in zip codes that also had no evidence of income misreporting. Differential experiences during the 2007–2010 bust period are even more striking. In high worse originator share zip codes, prices fell 32% compared to only 5% in zip codes with low worse originator share and no evidence of income misreporting. In short, zip codes without the originators most responsible for misreporting largely avoided the 2007–2010 house price crash. While this analysis falls well short of causal inference, the results with misreporting proxies are at least suggestive that misreporting played a significant role in enabling the house price cycle.

[Insert Figure 8 Here]

7 Conclusion

Although the forces behind the 2003–2006 housing price rise and the subsequent 2007–2010 collapse have been widely studied, there is surprisingly little consensus. We seek to make sense of the major competing academic views regarding forms of excess credit supply and speculation that potentially fueled housing demand by focusing on comparing empirical proxies. To put the competing explanations on equal footing, we construct all variables at the zip code level as of 2002. While most of the proxies have at least some correlation with house prices at the MSA level, two credit supply measures, subprime share and dubious origination practices, stand out for being systematically related to both house price increases during the boom and house price decreases during the bust. None of the speculative demand proxies are consistently related to within-MSA house price variation in both the boom and bust even after considering several possible constructs for extrapolative expectation.

The evidence indicates that credit supply amplified the house price cycle by relaxing lending standards and fueling demand. Subprime share and dubious origination practices are related to subsequent increases in non-owner occupancy and also transaction activity more generally, indicating that relaxed credit constraints in certain locales allowed individuals to speculate on housing. Inconsistent with lender or borrower extrapolative expectations driving future house price growth, we find that in areas of elastic housing supply, the credit channels are unrelated to 2003–2006 house price growth, but are related to increased speculation and transaction volume in 2003–2006 as well as larger price declines during the 2007–2010 house price bust. The effects of dubious origination practices are present in both agency and non-agency loans as well as subprime and wealthier zip codes, and are amplified in areas with potential income misreporting. Zip codes with low levels of fraudulent originators and income misreporting experienced minimal house price corrections.

Overall, the results present clear evidence that the distortive effects of excess credit through originators engaged in subprime and fraudulent practices facilitated speculation as well as general demand for housing, which led to a substantially distortive house price

boom and bust. Although we focus on the variables most widely used in the literature, use rich data, and perform substantial robustness tests, we recognize that our paper will likely not end debate and that future research may identify additional measures that may affect inferences. Our paper provides a systematic common framework for evaluating multiple credit supply and speculation proxies, which have previously not been compared to one another in the literature, and shows that the effects of credit supply through subprime and dubious origination practices were drivers for fueling speculation, transactions more generally, and house prices.

While many aspects of the financial crisis such as the players in the mortgage origination chain, the forces behind securitization, the banking panic, and the economic channels that interacted with house prices and amplified the financial crisis are beyond the scope of this paper, understanding the root causes of the housing crisis is central to understanding the financial crisis more generally. The forces behind the housing boom and bust are not merely of academic importance, but are also crucial for understanding what lessons to learn from the crisis and for guiding future policy. Our findings support the narrative proposed by many finance practitioners, the financial crisis inquiry commission, DOJ settlements, and substantial academic research indicating that deceptive fraudulent practices were widespread and systematic. While evidence of these practices was somewhat slow to emerge, it is perhaps not surprising that widespread systematic financing practices led to large distortions in house prices. As emphasized by Zingales (2015), we hope that future research can further understand and quantify costs associated with the 'rent-seeking' dimension of finance.

References

Adelino, Manuel, Antoinette Schoar, and Felipe Severino, 2016, Loan originations and defaults in the mortgage crisis: The role of the middle class, *Review of Financial Studies* 29, 1635–1670.

Adelino, Manuel, Antoinette Schoar, and Felipe Severino, 2018, Dynamics of housing debt in the recent boom and great recession, *NBER Macroeconomics Annual* 32, 265–311.

Albanesi, Stefania, Giacomo De Giorgi, and Jaromir Nosal, 2017, Credit growth and the financial crisis: A new narrative, Working Paper.

Ambrose, Brent W., James Conklin, and Jiro Yoshida, 2016, Credit rationing, income exaggeration, and adverse selection in the mortgage market, *Journal of Finance* 71, 2637–2686.

Anenberg, Elliot, Aurel Hizmo, Edward Kung, and Raven Molloy, 2017, Measuring mortgage credit availability: A frontier estimation approach, Working Paper.

Barlevy, Gadi, and Jonas Fisher, 2011, Mortgage choices and housing speculation.

Ben-David, Itzhak, 2011, Financial constraints and inflated home prices during the real estate boom, *American Economic Journal: Applied Economics* 3, 55–87.

Chinco, Alex, and Christopher Mayer, 2016, Misinformed speculators and mispricing in the housing market, *Review of Financial Studies* 29, 486–522.

Corbae, Dean, and Erwan Quintin, 2015, Leverage and the foreclosure crisis, *Journal of Political Economy* 123, 1–65.

DeFusco, Anthony A, Charles G Nathanson, and Eric Zwick, 2017, Speculative dynamics of prices and volume, Working Paper.

Demyanyk, Yuliya, and Otto Van Hemert, 2011, Understanding the subprime mortgage crisis, *Review of Financial Studies* 24, 1848–1880.

Di Maggio, Marco, and Amir Kermani, 2017, Credit-induced boom and bust, *Review of Financial Studies* 30, 3711–3758.

Elul, Ronel, and Sebastian G. Tilson, 2016, Owner occupancy fraud and mortgage performance, Working Paper.

Engel, Kathleen C., and Patricia A. McCoy, 2011, *The Subprime Virus: Reckless Credit, Regulatory Failure, and Next Steps* (Oxford University Press).

Favara, Giovanni, and Jean Imbs, 2015, Credit supply and the price of housing, *American Economic Review* 105, 958–92.

Fernandez, Carmen, Eduardo Ley, and Mark FJ Steel, 2001, Benchmark priors for Bayesian model averaging, *Journal of Econometrics* 100, 381–427.

Ferreira, Fernando, and Joseph Gyourko, 2015, A new look at the us foreclosure crisis: Panel data evidence of prime and subprime borrowers from 1997 to 2012, Working Paper.

Foote, Christopher L, Lara Loewenstein, and Paul S Willen, 2016, Cross-sectional patterns of mortgage debt during the housing boom: Evidence and implications, Working Paper.

Gao, Zhenyu, Michael Sockin, and Wei Xiong, 2017, Economic consequences of housing speculation, Working Paper.

Garmaise, Mark J., 2015, Borrower misreporting and loan performance, *Journal of Finance* 70, 449–484.

Gennaioli, Nicola, and Andrei Shleifer, 2018, *A Crisis of Beliefs: Investor Psychology and Financial Fragility* (Princeton University Press).

Glaeser, Edward L, Joshua D Gottlieb, and Joseph Gyourko, 2012, Can cheap credit explain the housing boom?, in *Housing and the financial crisis*, 301–359 (University of Chicago Press).

Glaeser, Edward L., Joseph Gyourko, and Albert Saiz, 2008, Housing supply and housing bubbles, *Journal of Urban Economics* 64, 198–217.

Glaeser, Edward L., and Charles G. Nathanson, 2017, An extrapolative model of house price dynamics, *Journal of Financial Economics* 126, 147–170.

Griffin, John M., and Gonzalo Maturana, 2016a, Did dubious origination practices distort house prices?, *Review of Financial Studies* 29, 1671–1708.

Griffin, John M., and Gonzalo Maturana, 2016b, Who facilitated misreporting in securitized loans?, *Review of Financial Studies* 29, 384–419.

Haughwout, Andrew, Donghoon Lee, Joseph Tracy, and Wilbert Van der Klaauw, 2011, Real estate investors, the leverage cycle, and the housing market crisis, FRB of New York Staff Report.

Herring, Richard J., and Susan Wachter, 2000, Real estate booms and banking busts, *Wharton Real Estate Review* 4.

Hubbard, R Glenn, and Christopher J. Mayer, 2009, The mortgage market meltdown and house prices, *The BE Journal of Economic Analysis & Policy* 9, Article 8.

Jiang, Wei, Ashlyn Aiko Nelson, and Edward Vytlacil, 2014, Liar's loan? Effects of origination channel and information falsification on mortgage delinquency, *Review of Economics and Statistics* 96, 1–18.

Kaplan, Greg, Kurt Mitman, and Giolvanni L. Violante, 2017, Non-durable consumption and housing net worth in the great recession: Evidence from easily accessible data, Working Paper.

Kermani, Amir, 2012, Cheap credit, collateral and the boom-bust cycle, Working Paper.

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

Keys, Benjamin J., Amit Seru, and Vikrant Vig, 2012, Lender screening and the role of securitization: Evidence from prime and subprime mortgage markets, *Review of Financial Studies* 25, 2071–2108.

Kolb, Robert, 2010, *Lessons from the Financial Crisis: Causes, Consequences, and Our Economic Future* (Wiley.).

Kruger, Samuel, and Gonzalo Maturana, 2018, Collateral misreporting in the RMBS market, Working Paper.

Levitin, Adam J., and Susan Wachter, 2012, Explaining the housing bubble, *Georgetown Law Journal* 100, 1177–1258.

Lewis, Michael, 2010, *The Big Short:Inside the Doomsday Machine* (W.W. Norton & Company).

Ley, Eduardo, and Mark FJ Steel, 2009, On the effect of prior assumptions in Bayesian model averaging with applications to growth regression, *Journal of Applied Econometrics* 24, 651–674.

Lowenstein, Roger, 2011, *The End of Wall Street* (Penguin Publishing Group).

Mayer, Christopher, 2011, Housing bubbles: A survey, *Annual Review of Economics* 3, 559–577.

Mayer, Christopher, Karen Pence, and Shane M. Sherlund, 2009, The rise in mortgage defaults, *Journal of Economic Perspectives* 23, 27–50.

McLean, Bethany, and Joseph Nocera, 2010, *All the Devils are Here: The Hidden History of the Financial Crisis* (Portfolio/Penguin Press).

Mian, Atif, and Amir Sufi, 2009, The consequences of mortgage credit expansion: Evidence from the US mortgage default crisis, *Quarterly Journal of Economics* 124, 1449–1496.

Mian, Atif, and Amir Sufi, 2017, Fraudulent income overstatment on mortgage applications during the credit expansion of 2002 to 2005, *Review of Financial Studies* 30, 1832–1864.

Mian, Atif, and Amir Sufi, 2018, Fueling a frenzy: Private label securitization and the housing cycle of 2000 to 2010, Working Paper.

Nadauld, Taylor D., and Shane M. Sherlund, 2013, The impact of securitization on the expansion of subprime credit, *Journal of Financial Economics* 107, 454–476.

Nathanson, Charles, and Eric Zwick, 2018, Arrested development: Theory and evidence of supply-side speculation in the housing market, *Journal of Finance.* Forthcoming.

Pavlov, Andrey, and Susan Wachter, 2011, Subprime lending and real estate prices, *Real Estate Economics* 39, 1–17.

Piskorski, Tomasz, Amit Seru, and James Witkin, 2015, Asset quality misrepresentation by financial intermediaries: Evidence from RMBS market, *Journal of Finance* 70, 2635–2678.

Purnanandam, Amiyatosh, 2011, Originate-to-distribute model and the subprime mortgage crisis, *Review of Financial Studies* 24, 1881–1915.

Rajan, Uday, Amit Seru, and Vikrant Vig, 2015, The failure of models that predict failure: Distance, incentives, and defaults, *Journal of Financial Economics* 115, 237–260.

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

Shiller, Robert J., 2006, Long-term perspectives on the current boom in home prices, *The Economists' Voice* 3.

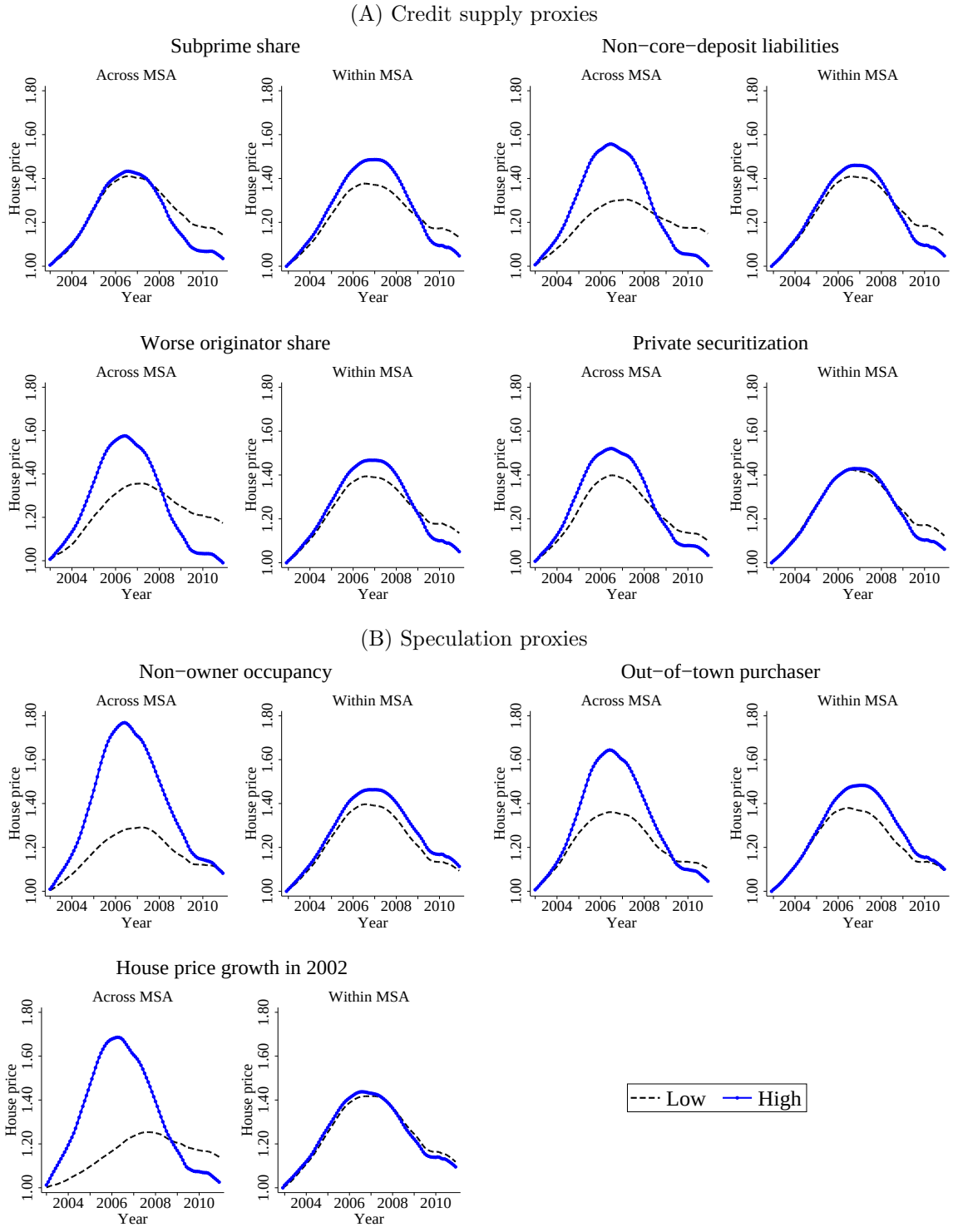
Shiller, Robert J., 2015, *Irrational exuberance: Revised and expanded third edition* (Princeton university press).

Varian, Hal R., 2014, Big data: New tricks for econometrics, *Journal of Economic Perspectives* 28, 3–28.

Vojtech, Cindy M., Benjamin Kay, and John C. Driscoll, 2016, The real consequences of bank mortgage lending standards, Working Paper.

Zingales, Luigi, 2015, Presidential address: Does finance benefit society?, *Journal of Finance* 70, 1327–1363.

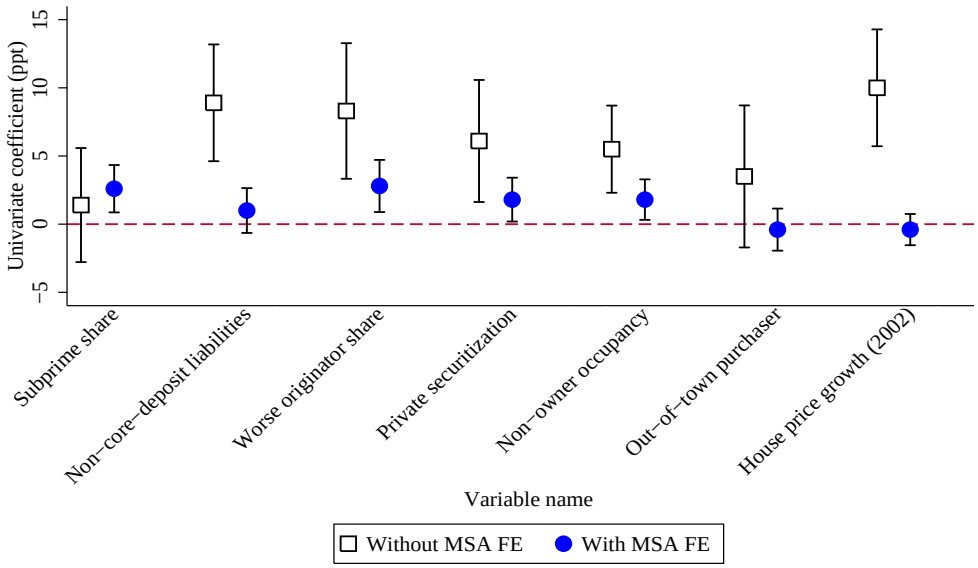
Figure 1. House price movements and 2002 proxies for credit supply and speculation



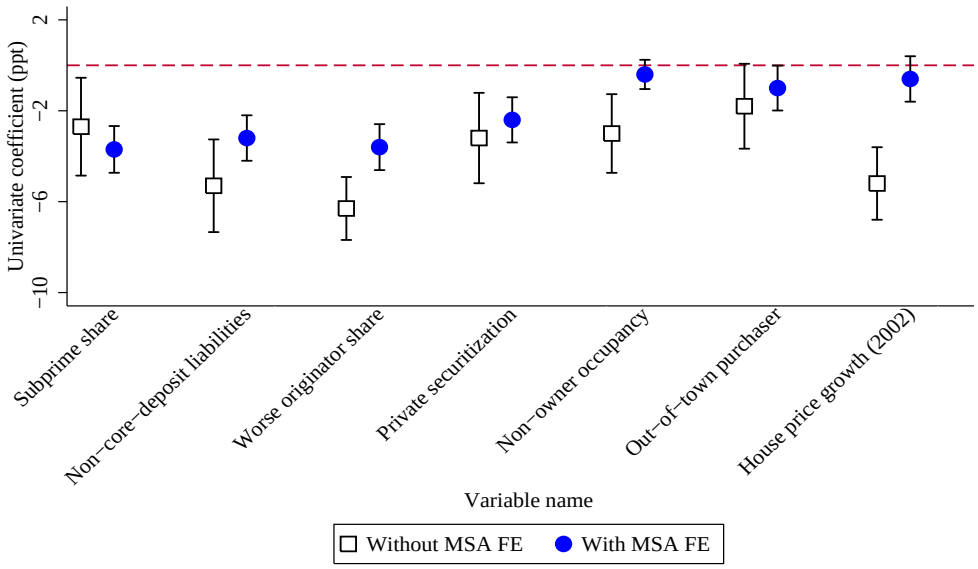
This figure shows the relation between house prices and proxies for exposure to credit supply (Panel A) and speculation (Panel B) based on MSA or zip-code level characteristics as of 2002. The plots on the left sort MSAs into quartiles based on the corresponding proxy. The plots on the right sort zip codes within a MSA into quartiles based on the corresponding proxy. The blue circles represent the average house price movement of the highest quartile whereas the solid black line represents the average house price movement of the lowest quartile. Variable definitions are the same as in Table 1.

Figure 2. Coefficient estimates of univariate house price growth regressions on 2002 proxies for credit supply and speculation

(A) 2003–2006 house price growth



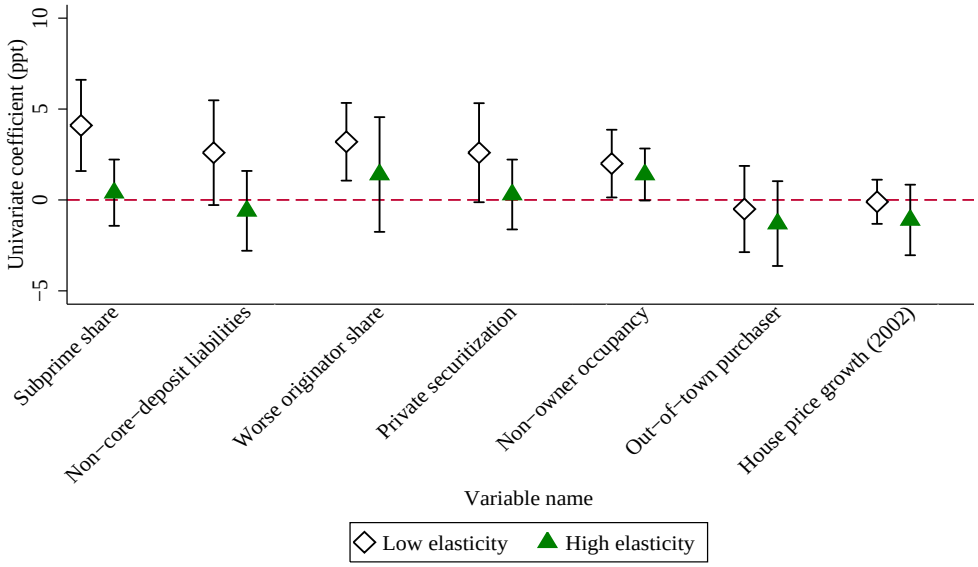
(B) 2007–2010 house price growth



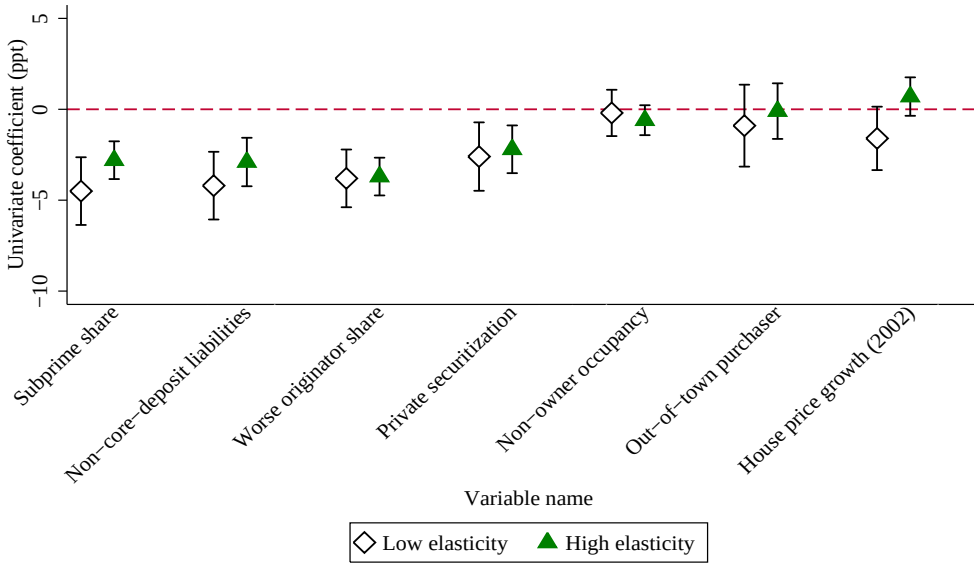
This figure shows the coefficients (with their corresponding 95% confidence intervals with standard errors clustered by MSA) for univariate regressions of zip-code level house price growth on proxies for exposure to credit supply and speculation based on zip-code level characteristics as of 2002. Panel A considers house price growth during 2003 to 2006 (i.e., price appreciation from December of 2002 to December of 2006). Panel B considers house price growth during 2007 to 2010 (i.e., price appreciation from December of 2006 to December of 2010). All proxies are standardized so that coefficients reflect the impact of changing the proxy by one standard deviation. Variable definitions are the same as in Table 1. All regressions control for population, housing units, and vacancy rates from 2000 census data, and 2002 average IRS income. Hollow squares represent coefficients of regressions without MSA fixed effects whereas solid circles represent coefficients of regressions with MSA fixed effects.

Figure 3. Coefficient estimates of univariate house price growth regressions on 2002 proxies for credit supply and speculation in subsamples based on housing supply elasticity

(A) 2003–2006 house price growth



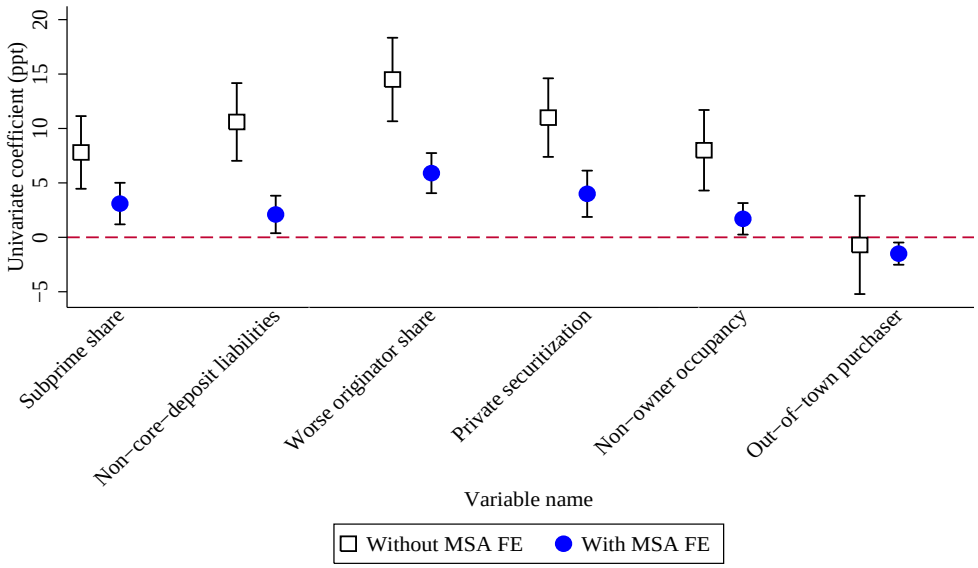
(B) 2007–2010 house price growth



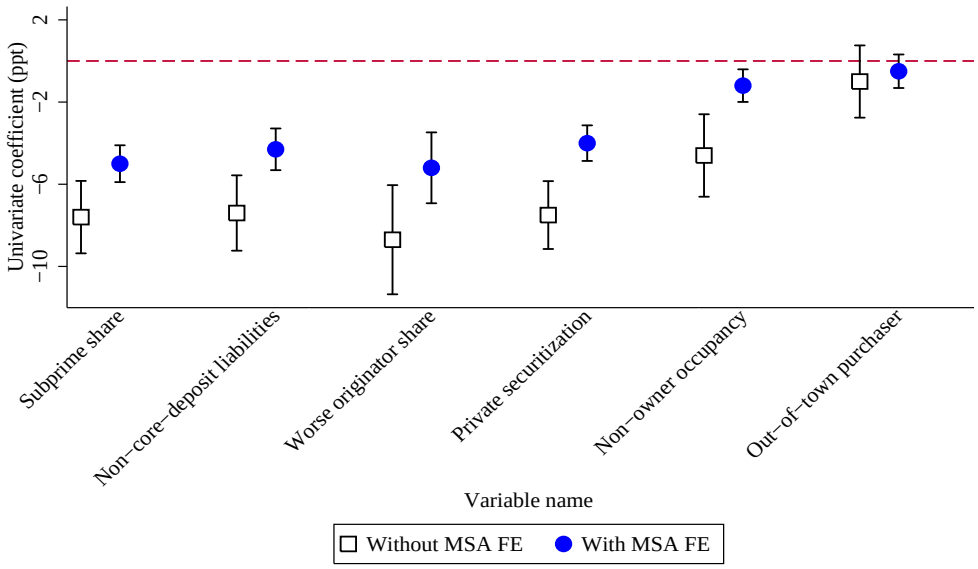
This figure shows the coefficients (with their corresponding 95% confidence intervals with standard errors clustered by MSA) for univariate regressions of zip-code level house price growth on proxies for exposure to credit supply and speculation based on zip-code level characteristics as of 2002. Panel A considers house price growth during 2003 to 2006 (i.e., price appreciation from December of 2002 to December of 2006). Panel B considers house price growth during 2007 to 2010 (i.e., price appreciation from December of 2006 to December of 2010). All proxies are standardized so that coefficients reflect the impact of changing the proxy by one standard deviation. Variable definitions are the same as in Table 1. All regressions control for population, housing units, and vacancy rates from 2000 census data, and 2002 average IRS income. Hollow diamonds represent coefficients of regressions where the sample corresponds of those zip codes in low housing supply elasticity areas (i.e., zip codes that have elasticity under the median level based on Saiz (2010)'s housing supply elasticity MSA-level measure) whereas solid triangles represent coefficients of regressions where the sample corresponds of those zip codes in high housing supply elasticity areas.

Figure 4. Coefficient estimates of univariate house price growth regressions on 2003–2006 proxies for credit supply and speculation

(A) 2003–2006 house price growth

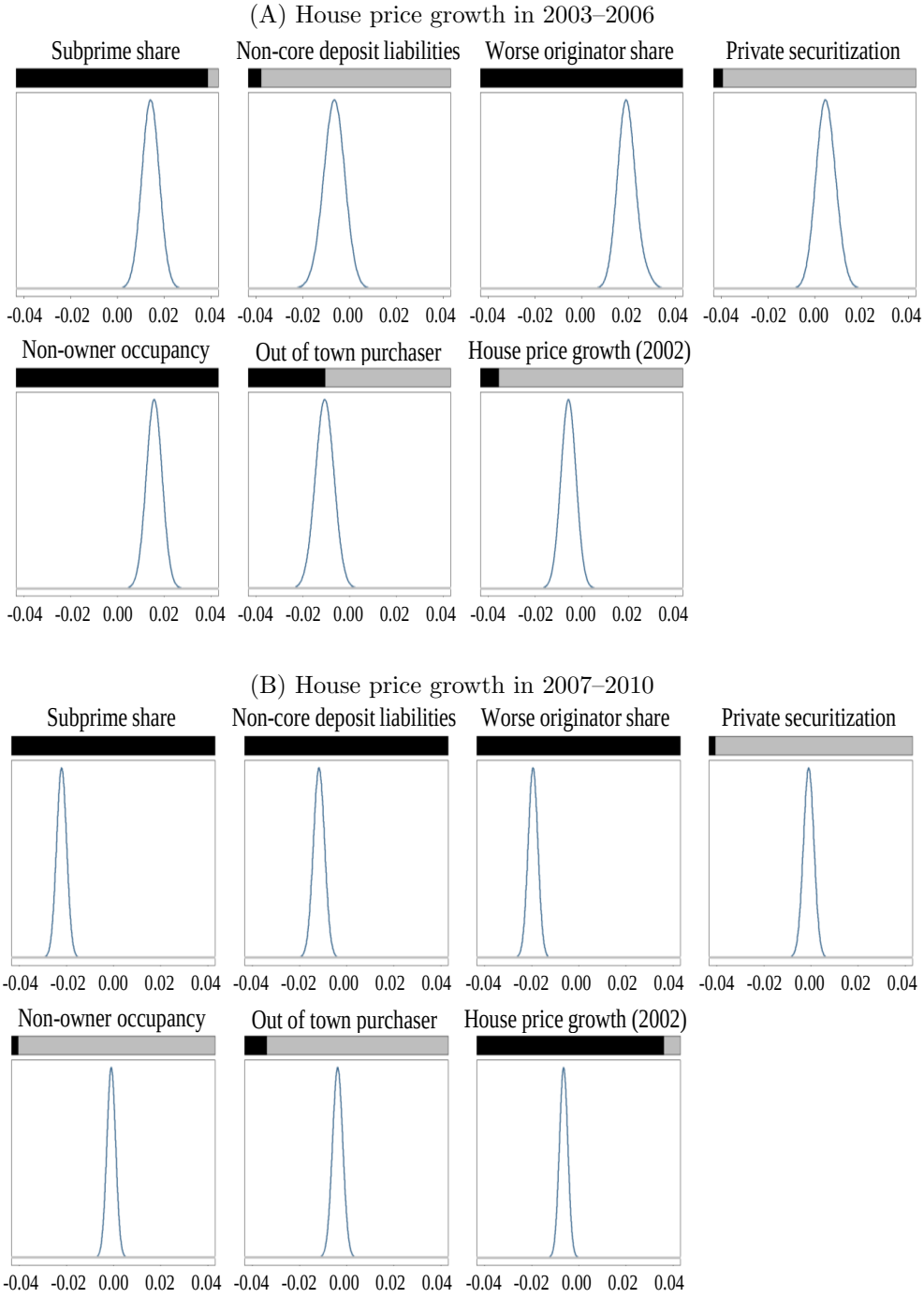


(B) 2007–2010 house price growth



This figure shows the coefficients (with their corresponding 95% confidence intervals with standard errors clustered by MSA) for univariate regressions of zip-code level house price growth on proxies for exposure to credit supply and speculation based on zip-code level characteristics from 2003 to 2006. Panel A considers house price growth during 2003 to 2006 (i.e., price appreciation from December of 2002 to December of 2006). Panel B considers house price growth during 2007 to 2010 (i.e., price appreciation from December of 2006 to December of 2010). All proxies are standardized so that coefficients reflect the impact of changing the proxy by one standard deviation. Variable definitions are the same as in Table 1. All regressions control for population, housing units, and vacancy rates from 2000 census data, and 2002 average IRS income. Hollow squares represent coefficients of regressions without MSA fixed effects whereas solid circles represent coefficients of regressions with MSA fixed effects.

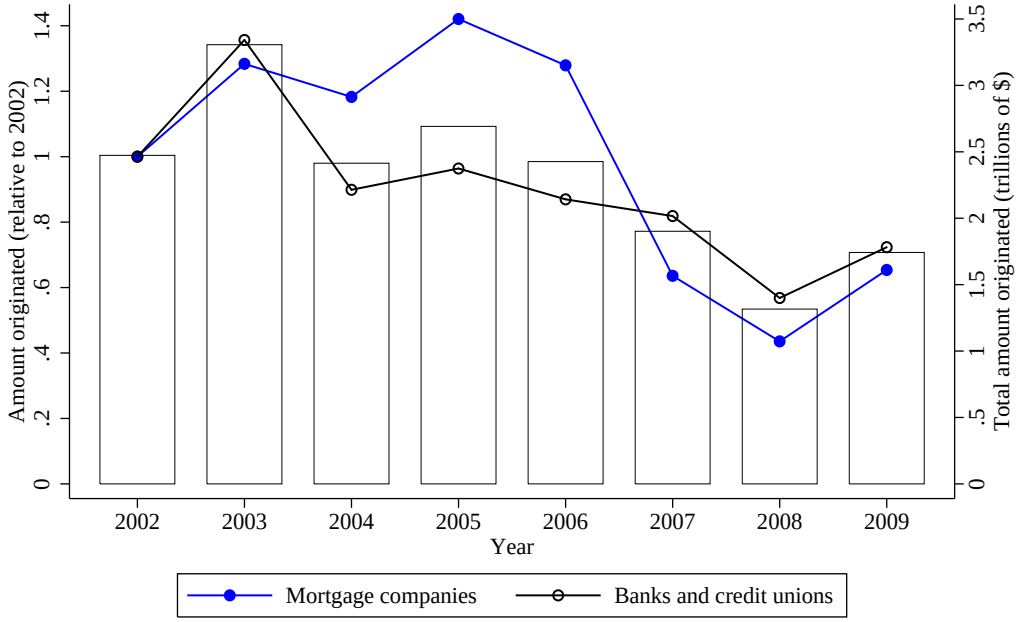
Figure 5. Posterior coefficient distributions



This figure plots posterior coefficient distributions from multivariate regressions of house price changes on 2002 proxies for credit supply and speculation using Bayesian model averaging. Panel A considers house price changes in 2003–2006, and Panel B considers house price changes in 2007–2010. The regressions include MSA fixed effects and control variables as described in Table 4. The black bar at the top of each figure plots posterior inclusion probability, and the lines in the plots represent posterior coefficient distributions conditional on the coefficient being included in the model.

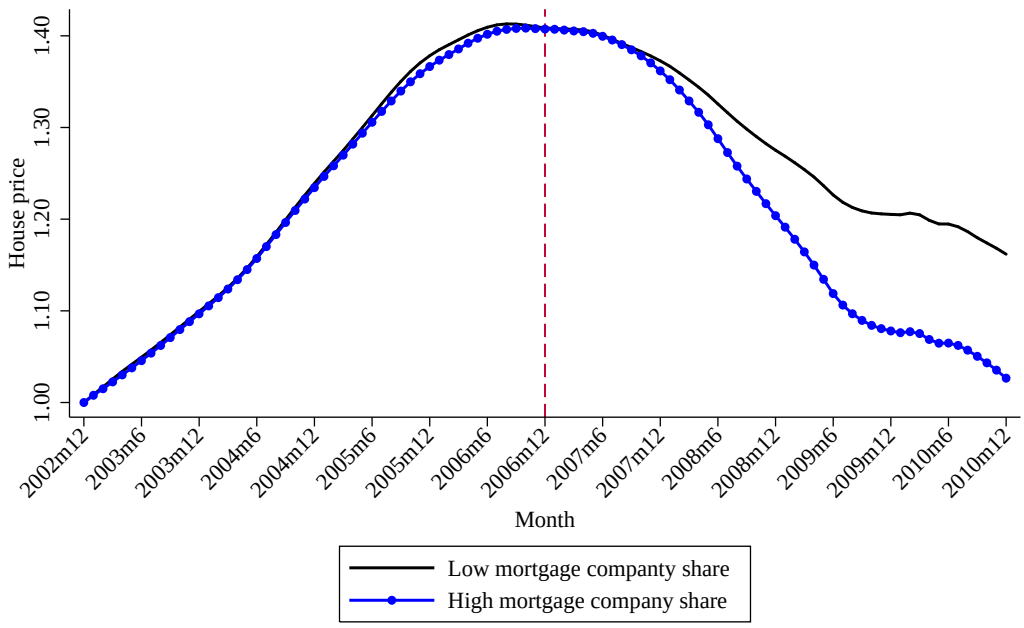
Figure 6. Mortgage volume and house price contraction

(A) Lending volume



This figure shows the relation between mortgage lending volume and house prices. Panel A shows total lending volume (represented by the bars), as well as lending volume by mortgage companies and banks and credit unions. The lending volume by mortgage companies and banks and credit unions is scaled by their corresponding lending volume in 2002. Panel B shows monthly house prices for high and low mortgage company share zip codes. Specifically, zip codes in the highest quartile of mortgage company share in 2006 are matched with zip codes in the lowest quartile of mortgage company share in 2006. Zip codes are required to be in the same MSA, to have a difference of no more than 1% in 2003–2006house price growth, and to have a difference of at least 15% in mortgage company share. Panel C plots coefficients from regressing house price changes on 2006 mortgage company market share with 95% confidence intervals based on standard errors clustered by MSA. Coefficients are from separate regressions of cumulative house price growth from the end of 2006 until the indicated quarter. Mortgage company market share is standardized so that coefficients reflect the impact of changing the variable by one standard deviation. Regressions are at the zip code level and control for 2003–2006 house price growth, population, housing units, and vacancy rates from 2000 census data, 2002 average IRS income, and MSA fixed effects.

(B) House price contraction



(C) Coefficient estimates of house price growth regressions on mortgage company share

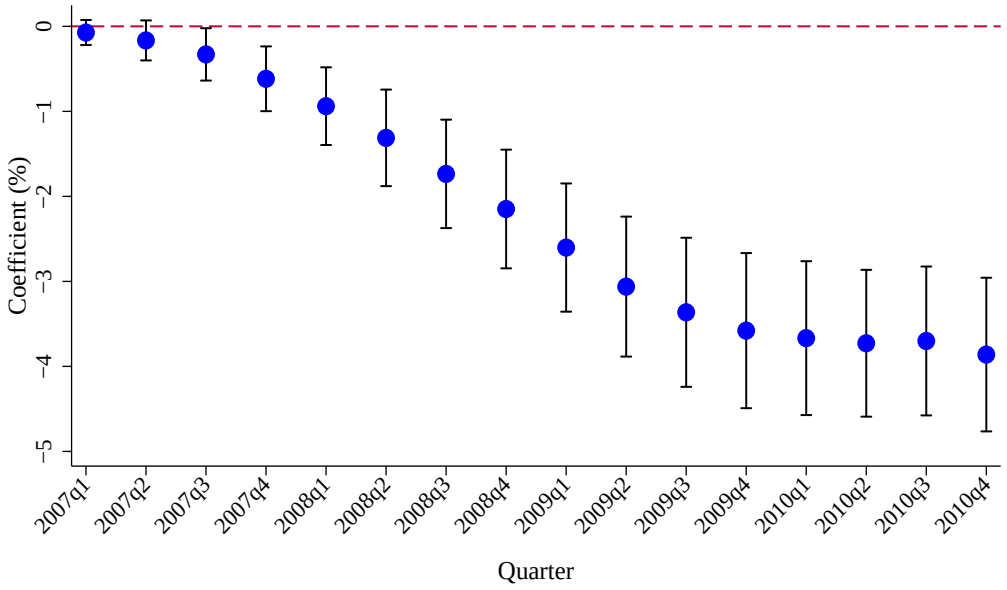
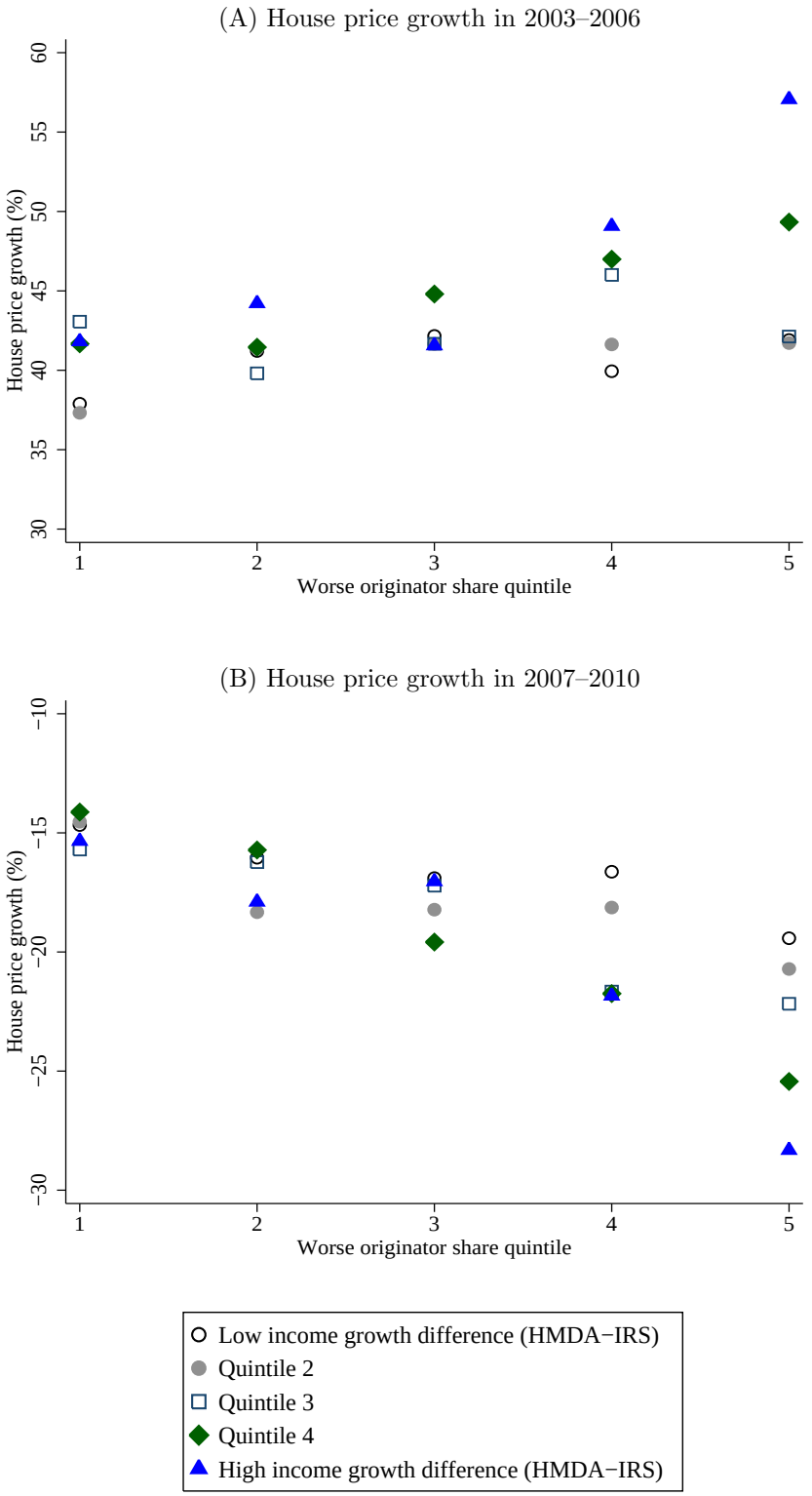
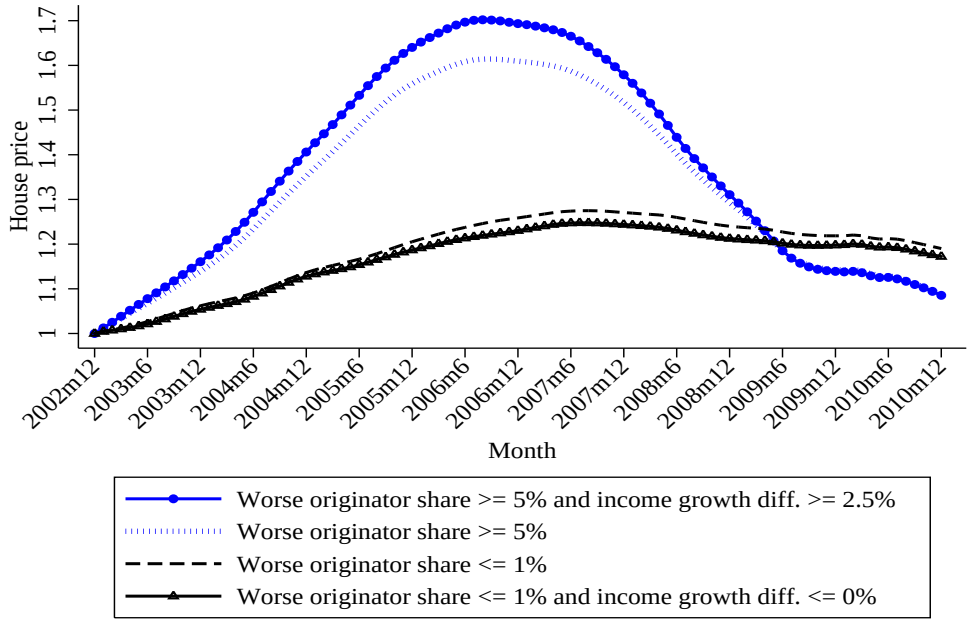


Figure 7. House price growth by worse originator share and income growth difference



This figure plots average house price growth in in 2003–2006 (panel A) and 2007–2010 (panel B) in zip codes classified by within-MSA quintiles based on 2003–2006 worse originator share and HMDA-IRS income growth differences.

Figure 8. House price movements in zip codes without misreporting



This figure plots average house price movement over time in zip codes classified based on worse originator share and HMDA-IRS income growth differences in 2003–2006.

Table 1. Data Summary

	Observations	Mean	Standard deviation		
			Overall	Between MSA	Within MSA
House price growth					
2003–2006	5,622	0.435	0.295	0.303	0.139
2007–2010	5,622	-0.189	0.156	0.168	0.078
2002 credit supply proxies					
Subprime share	3,725	0.094	0.058	0.039	0.050
Non-core-deposit liabilities	3,725	0.726	0.060	0.057	0.032
Worse originator share	3,725	0.017	0.016	0.012	0.012
Private securitization	3,725	0.291	0.083	0.075	0.051
2002 speculation proxies					
Non-owner occupancy	3,725	0.102	0.091	0.077	0.070
Out of town purchaser	3,725	0.061	0.079	0.095	0.044
House price growth (2002)	3,725	0.099	0.067	0.053	0.041
2003–2006 credit supply proxies					
Subprime share	5,622	0.134	0.070	0.042	0.062
Non-core-deposit liabilities	5,622	0.766	0.053	0.052	0.029
Worse originator share	5,622	0.038	0.026	0.019	0.019
Private securitization	5,622	0.420	0.087	0.078	0.054
2003–2006 speculation proxies					
Non-owner occupancy	5,622	0.129	0.100	0.085	0.074
Out of town purchaser	5,622	0.064	0.070	0.076	0.042

This table reports summary statistics for the main dependent variables and the proxies for exposure to credit supply and speculation. Zip-code level measures of each proxy are constructed both as of 2002 and contemporaneous with the 2003–2006 house price expansion. *Subprime share* is the fraction of mortgages originated by subprime lenders by zip code during the period of interest. The list of subprime lenders comes from the U.S. Department of Housing and Urban Development (HUD). *Non-core-deposit liabilities* is the weighted average (by market share) of lender non-core-deposit liability ratio (NCL) for all lenders with mortgage originations in the zip code. HMDA lenders are matched to call report data and NCL is calculated as $1-$*core deposits/total liabilities* aggregated to the bank holding company level. *Worse originator share* is the fraction of mortgages for originated by those originators in the highest second-lien misreporting tercile by zip code. The measure of second-lien misreporting is based on the comparison of what is reported by RMBS underwriters and whether second liens are actually present in county deed data. *Private securitization* is the fraction of mortgages sold to a private securitization trust or to a financial institution, an affiliated company, of an unspecified purchaser type. *Non-owner occupancy* is the fraction of mortgages for purchase associated with non-owner occupied properties. *Out-of-town purchaser* is the fraction of purchase transactions in a zip code where the buyer registered a mailing address outside the CSA where the property is located. *House price growth (2002)* is the return of the zip-code level Zillow house price index from December 2001 to December 2002.

Table 2. House price growth regressions on 2002 proxies for credit supply and speculation

Panel A: 2003–2006 house price growth

2002 zip code characteristics	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Credit supply							
Subprime share	0.026***						
	(2.935)						
Non-core-deposit liabilities		0.010					
		(1.192)					
Worse originator share			0.028***				
			(2.871)				
Private securitization				0.018**			
				(2.193)			
Speculation							
Non-owner occupancy					0.018**		
					(2.374)		
Out of town purchaser						-0.004	
						(-0.508)	
House price growth (2002)							-0.004
							(-0.685)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,725	3,725	3,725	3,725	3,725	3,725	3,725
R^2	0.847	0.844	0.847	0.845	0.845	0.844	0.844
Mean house price growth	0.463	0.463	0.463	0.463	0.463	0.463	0.463

This table reports coefficients for regressions of zip-code level house price growth on proxies for exposure to credit supply and speculation based on zip-code level characteristics as of 2002. Panel A considers house price growth during 2003 to 2006 (i.e., price appreciation from December of 2002 to December of 2006). Panel B considers house price growth during 2007 to 2010 (i.e., price appreciation from December of 2006 to December of 2010). All proxies are standardized so that coefficients reflect the impact of changing the proxy by one standard deviation. Variable definitions are the same as in Table 1. All regressions control for population, housing units, and vacancy rates from 2000 census data, 2002 average IRS income, and MSA fixed effects. t-statistics based on standard errors clustered by MSA are reported in parentheses. * indicates 10% significance, ** indicates 5% significance, and *** indicates 1% significance.

Panel B: 2007–2010 house price growth

2002 zip code characteristics	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Credit supply							
Subprime share	-0.037***						
	(-7.054)						
Non-core-deposit liabilities		-0.032***					
		(-6.270)					
Worse originator share			-0.036***				
			(-6.978)				
Private securitization				-0.024***			
				(-4.734)			
Speculation							
Non-owner occupancy					-0.004		
					(-1.218)		
Out of town purchaser						-0.010**	
						(-1.985)	
House price growth (2002)							-0.006
							(-1.176)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,725	3,725	3,725	3,725	3,725	3,725	3,725
R^2	0.826	0.812	0.823	0.810	0.802	0.802	0.802
Mean house price growth	-0.213	-0.213	-0.213	-0.213	-0.213	-0.213	-0.213

Table 3. House price growth regressions on lagged house price growth

	2003–2006 HPG			2003 HPG	2004 HPG	2005 HPG	2006 HPG
Lagged house price growth	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2000–2002 HPG	-0.023*						
	(-1.922)						
2002 HPG (5 mile radius)		0.001					
		(0.134)					
2000–2002 HPG (5 mile radius)			-0.019				
			(-1.363)				
2002 HPG				0.006**			
				(2.276)			
2003 HPG					-0.005**		
					(-2.086)		
2004 HPG						0.002	
						(0.391)	
2005 HPG							0.002
							(0.439)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,725	3,725	3,725	3,725	3,725	3,725	3,725
R^2	0.845	0.844	0.845	0.670	0.785	0.760	0.598
Mean house price growth	0.463	0.463	0.463	0.105	0.140	0.122	0.0266

This table reports coefficients for regressions of zip-code level house price growth on lagged house price growth. In columns (1) through (3) the dependent variable is house price growth during 2003 to 2006 (i.e., price appreciation from December of 2002 to December of 2006). The independent variables of interest are house price growth during 2000 to 2005 and refined versions of house price growth during 2002 and during 2000 to 2002 based on all zip codes within a 5 mile radius of the zip code being analyzed. In columns (4) to (7), yearly house price growth is regressed on house price growth lagged by one year. All independent variables are standardized so that coefficients reflect the impact of changing the independent variable by one standard deviation. All regressions control for population, housing units, and vacancy rates from 2000 census data, 2002 average IRS income, and MSA fixed effects. *t*-statistics based on standard errors clustered by MSA are reported in parentheses. * indicates 10% significance, ** indicates 5% significance, and *** indicates 1% significance.

Table 4. Multivariate house price growth regressions on 2002 proxies for credit supply and speculation

Panel A: 2003–2006 house price growth

	Regression coefficients and t-statistics				BMA posterior inclusion probability
2002 zip code characteristics	(1)	(2)	(3)	(4)	(5)
Credit supply					
Subprime share	0.015**		0.015*	0.014*	0.947
	(1.989)		(1.873)	(1.863)	
Non-core-deposit liabilities	-0.012		-0.009		0.062
	(-1.372)		(-1.032)		
Worse originator share	0.019**		0.019*	0.019*	1.000
	(1.993)		(1.938)	(1.971)	
Private securitization	0.008		0.006		0.046
	(0.747)		(0.628)		
Speculation					
Non-owner occupancy		0.019**	0.016**	0.015**	0.998
		(2.389)	(1.999)	(2.052)	
Out of town purchaser		-0.007	-0.011		0.380
		(-0.805)	(-1.280)		
House price growth (2002)		-0.005	-0.005		0.090
		(-0.892)	(-0.900)		
Control variables	Yes	Yes	Yes	Yes	Yes
MSA fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	3,725	3,725	3,725	3,725	3,725
R^2	0.848	0.845	0.850	0.849	
Mean house price growth	0.463	0.463	0.463	0.463	0.463
Credit supply F-test (p-value)	0.013**		0.029**		
Speculation F-test (p-value)		0.131	0.195		

This table reports coefficients for regressions of zip-code level house price growth on proxies for exposure to credit supply and speculation based on zip-code level characteristics as of 2002. Columns (1) through (4) show coefficients and t-statistics and column (5) reports posterior inclusion probabilities for each proxy based on Bayesian model averaging. Panel A considers house price growth during 2003 to 2006 (i.e., price appreciation from December of 2002 to December of 2006). Panel B considers house price growth during 2007 to 2010 (i.e., price appreciation from December of 2006 to December of 2010). All proxies are standardized so that coefficients reflect the impact of changing the proxy by one standard deviation. Variable definitions are the same as in Table 1. All regressions control for population, housing units, and vacancy rates from 2000 census data, 2002 average IRS income, and MSA fixed effects. t-statistics based on standard errors clustered by MSA are reported in parentheses. * indicates 10% significance, ** indicates 5% significance, and *** indicates 1% significance.

Panel B: 2007–2010 house price growth

	Regression coefficients and t-statistics				BMA posterior inclusion probability
2002 zip code characteristics	(1)	(2)	(3)	(4)	(5)
Credit supply					
Subprime share	-0.021***		-0.021***	-0.022***	1.000
	(-4.178)		(-4.121)	(-4.558)	
Non-core-deposit liabilities	-0.012***		-0.011***	-0.012***	1.000
	(-2.778)		(-2.714)	(-3.034)	
Worse originator share	-0.020***		-0.019***	-0.019***	1.000
	(-3.427)		(-3.421)	(-3.447)	
Private securitization	-0.000		-0.001		0.031
	(-0.050)		(-0.164)		
Speculation					
Non-owner occupancy		-0.003	-0.001		0.032
		(-0.837)	(-0.192)		
Out of town purchaser		-0.010**	-0.004		0.107
		(-2.088)	(-0.948)		
House price growth (2002)		-0.007	-0.007	-0.006	0.981
		(-1.199)	(-1.383)	(-1.322)	
Control variables	Yes	Yes	Yes	Yes	Yes
MSA fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	3,725	3,725	3,725	3,725	3,725
R^2	0.831	0.803	0.832	0.832	
Mean house price growth	-0.213	-0.213	-0.213	-0.213	-0.213
Credit supply F-test (p-value)	0.000***		0.000***		
Speculation F-test (p-value)		0.033**	0.184		

Table 5. Housing demand impact of credit supply proxies

	2003–2006		2003–2006 house purchase volume					
	non-owner occupancy		total		short term (< 1 year)		long term (> 1 year)	
2002 zip code characteristics	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Subprime share	0.005***		0.004***		0.003***		0.001**	
	(3.118)		(7.106)		(10.308)		(2.468)	
Worse originator share		0.004**		0.003***		0.002***		0.002***
		(2.216)		(6.843)		(4.239)		(4.095)
Non-owner occupancy	0.073***	0.074***						
	(26.107)	(26.190)						
House purchase volume			0.021***	0.021***	0.005***	0.005***	0.016***	0.015***
			(15.057)	(15.155)	(11.590)	(10.804)	(13.887)	(13.899)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,725	3,725	3,725	3,725	3,725	3,725	3,725	3,725
R^2	0.898	0.897	0.867	0.865	0.844	0.823	0.842	0.843
Dependent variable mean	0.134	0.134	0.0838	0.0838	0.0129	0.0129	0.0708	0.0708

This table reports coefficients for regressions of zip-code level non-owner occupancy and house purchase volume on proxies for exposure to credit supply based on zip-code level characteristics as of 2002. Non-owner occupancy is the fraction of mortgages for purchase associated with non-owner occupied properties. House purchase volume is the number of purchase transactions per zip code divided by the existing number of residential properties in the zip code. Short term volume is defined as transaction in which the property was last purchased within the past year. Long term volume consists of all other transactions. All proxies and control variables are standardized so that coefficients reflect the impact of changing the proxy by one standard deviation. Variable definitions are the same as in Table 1. All regressions control for population, housing units, and vacancy rates from 2000 census data, 2002 average IRS income, and MSA fixed effects. t-statistics based on standard errors clustered by MSA are reported in parentheses. * indicates 10% significance, ** indicates 5% significance, and *** indicates 1% significance.

Table 6. Housing demand in areas with high housing supply elasticity

	2003–2006		2003–2006 house purchase volume					
	non-owner occupancy		total		short term (< 1 year)		long term (> 1 year)	
2002 zip code characteristics	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Subprime share	0.008***		0.002***		0.003***		-0.000	
	(3.201)		(3.226)		(9.289)		(-0.606)	
Worse originator share		0.007***		0.003***		0.002***		0.001
		(2.863)		(3.066)		(5.259)		(1.173)
Non-owner occupancy	0.076***	0.077***						
	(21.350)	(22.890)						
House purchase volume			0.022***	0.021***	0.005***	0.005***	0.017***	0.017***
			(8.388)	(8.513)	(6.168)	(6.133)	(8.995)	(9.078)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,636	1,636	1,636	1,636	1,636	1,636	1,636	1,636
R^2	0.859	0.857	0.862	0.863	0.821	0.806	0.846	0.846
Dependent variable mean	0.0783	0.0783	0.0783	0.0783	0.0120	0.0120	0.0662	0.0662

This table reports coefficients for regressions of zip-code level non-owner occupancy and house purchase volume on proxies for exposure to credit supply based on zip-code level characteristics as of 2002. Only zip codes in high housing supply elasticity areas (i.e., zip codes that have elasticity above the median level based on Saiz (2010)'s housing supply elasticity MSA-level measure) are included. Non-owner occupancy is the fraction of mortgages for purchase associated with non-owner occupied properties. House purchase volume is the number of purchase transactions per zip code divided by the existing number of residential properties in the zip code. Short term volume is defined as transaction in which the property was last purchased within the past year. Long term volume consists of all other transactions. All proxies and control variables are standardized so that coefficients reflect the impact of changing the proxy by one standard deviation. Variable definitions are the same as in Table 1. All regressions control for population, housing units, and vacancy rates from 2000 census data, 2002 average IRS income, and MSA fixed effects. t-statistics based on standard errors clustered by MSA are reported in parentheses. * indicates 10% significance, ** indicates 5% significance, and *** indicates 1% significance.

Table 7. Credit expansion and worse originator market share

	Full sample		Low subprime zip codes	
2002 zip code characteristics	(1) Mortgage growth	(2) IRS income growth	(3) Mortgage growth	(4) IRS income growth
Worse originator share	0.083***	-0.003***	0.067***	-0.001
	(7.165)	(-3.776)	(3.064)	(-1.165)
Control variables	Yes	Yes	Yes	Yes
MSA fixed effects	Yes	Yes	Yes	Yes
Observations	3,725	3,725	1,863	1,863
R^2	0.592	0.559	0.582	0.570
Dependent variable mean	0.145	0.0398	0.0960	0.0381

This table reports coefficients for regressions of zip-code level mortgage origination growth (columns (1) and (2)) and income growth (columns (3) and (4)) on worse originator share in 2002. Mortgage origination growth is average annual mortgage origination dollar volume in 2003 to 2006 compared to 2002. Income growth is 2002–2006 growth in IRS income data. Columns (1) and (2) report results for the full sample. Columns (3) and (4) report results restricted to zip codes with 2002 subprime lender market shares in the bottom quartile. Worse originator share is standardized so that coefficients reflect the impact of changing worse originator share by one standard deviation. Regressions control for population, housing units, and vacancy rates from 2000 census data, 2002 average IRS income, and MSA fixed effects. t-statistics based on standard errors clustered by MSA are reported in parentheses. * indicates 10% significance, ** indicates 5% significance, and *** indicates 1% significance.

Table 8. Loan performance

	(1)	(2)	(3)
Worse originator share	0.007**	0.007**	0.007**
	(2.55)	(2.55)	(2.31)
Worse originator share			0.003***
×Unidentified originator			(4.62)
Control variables	Yes	Yes	Yes
Lender fixed effects	No	Yes	Yes
MSA×origination quarter fixed effects	Yes	Yes	Yes
Observations	5,546,136	5,546,136	5,546,136
R^2	0.105	0.105	0.106
Mean	0.077	0.077	0.077

This table reports results from regressions of loan delinquency on zip code level market share of worse originators, detailed loan characteristics, originator fixed effects (as indicated), and MSA×origination quarter fixed effects. Observations represent individual loans originated in 2003 to 2006, and the dependent variable is an indicator for whether the loan became 90 or more days delinquent at any point between origination and September 2012. The sample consists of loans that were securitized by Fannie Mae. Unidentified originators are originators that represent less than 1% of volume within a given acquisition quarter as represented by the original unpaid principal balance in the Fannie Mae data, which are grouped together in the data instead of being individually identified. For originator fixed effects, these originators are grouped together. All regressions control for log loan size, loan-to-value ratio, credit score, interest rate, a 0/1 indicator for purchase, and a 0/1 indicator for owner-occupied home. t-statistics based on standard errors clustered by MSA are reported in parentheses. * indicates 10% significance, ** indicates 5% significance, and *** indicates 1% significance.

Table 9. Worse originator market share and HMDA–IRS income growth difference

	2002–2006 income growth difference (HMDA - IRS)	2003–2006 house price growth			2007–2010 house price growth		
2003–2006 zip code characteristics	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Worse originator share	0.011***		0.056***	0.030**		-0.051***	-0.049***
	(7.161)		(5.604)	(2.286)		(-5.894)	(-6.893)
Income growth difference (HMDA - IRS)		0.017***	0.008*	0.011***	-0.012***	-0.004	-0.004
		(2.809)	(1.851)	(2.685)	(-4.539)	(-1.634)	(-1.567)
Worse originator share ×Income growth difference (HMDA - IRS)				0.030***			-0.002
				(4.929)			(-0.607)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,622	5,622	5,622	5,622	5,622	5,622	5,622
R^2	0.458	0.820	0.831	0.838	0.787	0.819	0.819
Dependent variable mean	0.00586	0.435	0.435	0.435	-0.189	-0.189	-0.189

This table reports coefficients for regressions of zip-code level income growth differences and house price growth on worse originator market share and the difference between income growth in HMDA and IRS data in 2003 to 2006. Column (1) reports results for a regression of income growth differences on worse originator share. Columns (2) and (4) consider house price growth during 2003 to 2006 (i.e., price appreciation from December of 2002 to December of 2006). Columns (5) and (7) consider house price growth during 2007 to 2010 (i.e., price appreciation from December of 2006 to December of 2010). Explanatory variables are standardized so that coefficients reflect the impact of changing the proxy by one standard deviation. All regressions control for population, housing units, and vacancy rates from 2000 census data, 2002 average IRS income, and MSA fixed effects. t-statistics based on standard errors clustered by MSA are reported in parentheses. * indicates 10% significance, ** indicates 5% significance, and *** indicates 1% significance.