Fueling a Frenzy: Private Label Securitization and the Housing Cycle of 2000 to 2010

Atif Mian
Princeton University and NBER

Amir Sufi
University of Chicago Booth School of Business and NBER

March 2018

Abstract

Geographic areas more exposed to the 2003 acceleration of the private label mortgage securitization (PLS) market witness a sudden and large increase in mortgage originations and transaction volume from 2003 to 2006. These areas experience significant relative growth in the number of individuals with many mortgages, highlighting the importance of a small group of individuals in driving the rise in volume. House prices and construction activity grow substantially more in these areas. Cities such as Las Vegas and Phoenix with high exposure to the PLS market are significantly more likely to experience a simultaneous large increase in both house price growth and construction activity during the housing boom. These cities see a painful bust, with house prices and construction activity falling below pre-2003 levels. The results are inconsistent with the view that a general rise in housing market optimism can explain house price growth in cities most exposed to the PLS market; to the contrary, higher house price growth driven by the acceleration of the PLS market boosts the share of individuals saying it is a bad time to buy a house because prices are too high. Overall, the results suggest that credit supply expansion fueled by the PLS market allowed a small group of individuals to have large effects on the housing market.

*This research was supported by funding from the Julis Rabinowitz Center For Public Policy and Finance at Princeton, and the Fama-Miller Center and Initiative on Global Markets at Chicago Booth. We are grateful to Tomasz Piskorski and Amit Seru for feedback and help with data questions. Tu Cao and Michael Varley provided excellent research assistance. Mian: (609) 258 6718, atif@princeton.edu; Sufi: (773) 702 6148, amir.sufi@chicagobooth.edu

1 Introduction

The United States experienced an amplified housing cycle from 2000 to 2010, and a consensus has emerged that this cycle was an important factor precipitating the Great Recession. However, there remains disagreement on the underlying factors responsible for the housing cycle. A variety of explanations have been put forth by researchers, including shifts in mortgage credit supply, changes in housing market optimism, and speculation by investors.

This study presents evidence that the amplitude of the housing cycle is explained in part by the acceleration and subsequent collapse of the private-label mortgage backed securitization market (the PLS market hereafter).[1] Figure 1 shows the acceleration and subsequent collapse of the PLS market. Following the large mortgage refinancing wave of 2000 to 2002, there was a significant rise in the fraction of originated mortgages sold into the PLS market. By 2006, almost half of new dollars originated were sold into the PLS market.

The existing literature suggests that the acceleration of the PLS market was a credit supply-side phenomenon associated with a large increase in mortgage quantity and a significant decline in mortgage interest spreads (e.g., Levitin and Wachter (2013); Justiniano et al. (2017)). However, the rise in the PLS market could have simply reflected changes in household demand-side factors such as a rise in income or a change in expectations of future house price growth. Analysis of aggregate data alone cannot easily distinguish whether the PLS market had an effect on the housing market independent of shocks to household income or beliefs.

The empirical strategy used in this study isolates plausibly exogenous variation in mortgage origination growth across U.S. geographical areas due to the acceleration of the PLS market in 2003, and it demonstrates a significant independent effect of the acceleration of the PLS market on the housing cycle from 2003 to 2010. The strategy begins with cross-sectional variation across financial institutions in reliance on non-core deposits to fund loans.

[1] The PLS market during the 2000s included subprime mortgages, but subprime mortgage originations made up no more than 40% of this market in any year from 2000 to 2006. It is important to emphasize from the outset that this market is broader than the subprime segment.

In particular, the acceleration of the PLS market after 2002 disproportionately reduced the cost of originating mortgages for financial institutions that traditionally relied on non-core deposit liability financing.[2] The extreme example of such financial institutions is the group of non-bank mortgage lenders, such as Ameriquest Mortgage Company, that rely entirely on non-core deposit financing in the mortgage origination process.

The bank-level results show that even regulated deposit-taking financial institutions that traditionally relied more on non-core deposit liability financing prior to the rise of the PLS market witnessed a sudden relative expansion in mortgage lending starting in 2003. In particular, among deposit-taking financial institutions, the non-core deposit liabilities to total liabilities ratio (the NCL ratio hereafter) as of 2002 predicts a large and statistically robust increase in mortgage originations from 2002 to 2006. Further, this expansion in mortgage lending by high NCL ratio banks is driven by mortgages sold into the PLS market. The analysis combines mortgage lenders and traditional banks relying heavily on non-core deposit financing into one group (which we call "High NCL lenders"), and it shows that mortgage originations by these financial institutions see strong relative growth beginning exactly with the acceleration of the PLS market in 2003.

High NCL lenders as of 2002 were not equally distributed across the country, a fact the empirical strategy exploits to generate variation across geographical areas in exposure to the acceleration of the PLS market in 2003. In particular, the Home Mortgage Disclosure Act (HMDA) data is used to calculate the average NCL ratio for lenders originating mortgages in a given zip code, where the average NCL ratio is weighted by the amount of mortgage originations by the lender in the zip code in 2002.[3]

Areas more exposed to high NCL lenders as of 2002 see strong relative growth in mortgage

[2]Core deposits are defined by the FDIC as " the sum of demand deposits, all NOW and automatic transfer service (ATS) accounts, money market deposit accounts (MMDAs), other savings deposits, and time deposits under $100,000." As the FDIC writes, "Core deposits, as an analytical and supervisory tool, are intended to include those deposits that are stable and lower cost and that reprice more slowly than other deposits when interest rates rise" (FDIC 2011).

[3]To the best of our knowledge, there are only two other studies that specify an exogenous source of variation across U.S. geographical areas to study the housing boom from 2000 to 2007: Di Maggio and Kermani (2017) and Gao et al. (2016).

amounts originated beginning in late 2003 and ending in 2006, a time period that corresponds exactly to the acceleration of the PLS market. High NCL share zip codes see stronger origination growth for both new home-purchase and refinancing mortgage amounts. Further, this growth is especially strong for mortgages that are subsequently sold into the PLS market.

The sudden and large increase in originated mortgage amounts sparks a trading frenzy in high NCL share zip codes. We use two measures of transaction volume: the number of new purchase mortgages and a measure of volume from deed records collected by DataQuick. For both measures, the results show a significant increase in volume from 2002 to 2006 in zip codes most exposed to high NCL lenders as of 2002. In terms of magnitudes, a one standard deviation increase in NCL exposure leads 10 to 20% increase in transaction volume.

A substantial body of research suggests that investors or "flippers" were crucial in explaining the housing boom from 2000 to 2007.[4] The empirical analysis in this study uses data from TransUnion to measure the number of individuals in a zip code with many mortgages. Zip codes most exposed to the acceleration of the PLS market in 2003 see substantial relative growth in the number of individuals with five or more mortgages. In contrast, these zip codes see no relative growth in the number of individuals with one or more mortgages. We interpret this result as showing how the acceleration of the PLS market significantly increased the buying power of a small group of investors.

The zip code level evidence suggests that the acceleration of the PLS market in 2003 represented a positive credit supply shock from 2003 to 2006, and this shock led to a large increase in transaction volume. This shock also leads to significant relative house price and construction growth in high PLS share areas from 2002 to 2006. In terms of magnitudes, a one standard deviation increase in the NCL ratio leads to an increase in house price growth of 6 percentage points (15 percentage points in the most inelastic housing supply areas), and a rise in construction of 1% of the housing stock.

One of the puzzling aspects of the housing boom from 2002 to 2006 is the presence of cities

[4]See, e.g., DeFusco et al. (2017); Chinco and Mayer (2015); Bhutta (2015); Haughwout et al. (2011).

with elastic housing supply that simultaneously experience large house price and construction growth (e.g., Glaeser et al. (2008), Davidoff (2013), Nathanson and Zwick (2017)). As shown Glaeser et al. (2008), these anomalous MSAs are difficult to explain in standard models with rational beliefs; if housing supply is elastic and there are a finite number of borrowers, then it is difficult to justify large increases in house prices.

The findings presented here suggest that the rise of the PLS market is a crucial factor in explaining this puzzle. Exposure to the acceleration of the PLS market measured as of 2002 robustly predicts whether a city experiences a simultaneous increase in construction and house prices during the boom. Further, cities with a large NCL share experience an "over-correction" to both house price growth and construction growth during the bust. By 2010, house prices and construction activity are below the 2002 level in high NCL share cities. Glaeser et al. (2008) make exactly this prediction in a model with irrational beliefs: excessive building in elastic housing supply cities during the boom will lead to a steep decline in prices and construction when prices eventually fall.[5]

Survey evidence on housing market optimism supports the view that the PLS market affected house prices by giving a small group of individuals significant buying power. Using individual responses to the Michigan Survey of Consumers (Piazzesi and Schneider (2009)), the analysis does not find evidence that individuals living in MSAs most exposed to high NCL lenders experienced a statistically significant relative change in attitudes toward home buying prior to the rise of the PLS market. In other words, we do not find evidence that a local "optimism shock" started the lending boom in high NCL lender share MSAs.

In contrast, MSAs with higher house price growth driven by the acceleration of the PLS market actually experience a relative increase in the fraction of individuals saying that now is a bad time to buy a home. This is driven by individuals saying that now is a bad time to

[5]Nathanson and Zwick (2017) construct a model with supply-side speculators that can explain this pattern. The findings presented here suggest that the presence of the PLS market was an important factor in predicting where such speculation arose. These findings are also consistent with a suggestion in Davidoff (2013) who notes that the presence of certain banks as of 2001 in many of these anomalous markets may explain house price volatility.

buy a home because prices are too high or prices will fall. Our interpretation of this result is that an expansion of credit supply through the PLS market increased the buying power of a small group of individuals, thereby boosting house prices. The average individual in the MSA responded to higher house prices by becoming more pessimistic about buying a house. The findings in this setting are inconsistent with the view that a uniform increase in housing market optimism was a driver of house price growth from 2002 to 2006. Instead, in combination with the findings on investors, the results are more supportive of models in which credit supply expansion boosts asset prices by giving substantial buying power to a small group of individuals (e.g., Geanakoplos (2010), Simsek (2013), Burnside et al. (2016)).

A concern with the findings is omitted variable bias: perhaps high NCL share zip codes experienced other shocks from 2002 to 2006 that can explain the results. The allocation of high NCL share lenders across the country prior to 2003 is not random. High NCL lenders concentrate their mortgage lending in deposit-poor MSAs that have higher average house prices and more inelastic housing supply conditions. Within these MSAs, high NCL lenders focus on lower income, lower credit score, and higher minority share zip codes.

There are three findings that mitigate the concern that alternative shocks are responsible for the results. First, specifications conducted at the lender-MSA and lender-zip code level show that high NCL lenders expand mortgage originations more than low NCL lenders even within the same narrow geographical area. For example, the relative growth in originated mortgage amounts for high versus low NCL lenders is large and statistically significant at the 1% level even with the inclusion of zip code fixed effects. This implies that unobservable characteristics or shocks at the zip code level cannot explain the relative growth in originated mortgage amounts by high NCL lenders during the housing boom. Further, the effect of the NCL lender share on house prices and volume is qualitatively similar when using only within-MSA zip-code level variation in the NCL lender share.

Second, the timing of the relative growth in mortgage originations in high NCL MSAs matches exactly the acceleration of the PLS market in mid-2003. There is no evidence of differential pre-trends in mortgage originations prior to 2003. Further, we use county-by-month HMDA data from the Federal Reserve to show that the relative increase in mortgage originations in high NCL share MSAs begins in September 2003, just as the PLS market accelerates in response to the decline in refinancing originations. The relative rise in mortgage originations in high NCL share MSAs corresponds exactly with the dramatic decline in the PLS mortgage interest spread to Treasury rate in August and September of 2003, called "the mortgage rate conundrum" by Justiniano et al. (2017). The relative growth in volume, house prices, and construction in high NCL areas also starts in 2003 with the acceleration of the PLS market.[6] Third, we examine the main alternative hypothesis put forth in the literature: that a change in beliefs about house prices can explain the rise in house prices. As already mentioned, we are unable to find evidence that optimism about house price growth changed differentially in high NCL share MSAs before 2003.

There is a large body of research on the PLS market and the housing boom and bust from 2000 to 2010. We first present our main results, and then we discuss how our findings are related to this large body of research in Section 7.

2 Data and Summary Statistics

2.1 Data

The main data sets used in this study are at the lender, MSA, and zip code level. We begin with the Home Mortgage Disclosure Act (HMDA) data which records the universe of mortgage originations for mortgage originators that have an office within metropolitan statistical areas (MSAs).[7] We identify each mortgage originator in the HMDA data, and we classify them as either a "bank" or a "non-bank" based on whether they are regulated by

[6]The results are almost identical if we use a measure of exposure to high NCL lenders as of 1998 instead of 2002. These results are reported in the appendix.

[7]See guidelines for HMDA issued by the Federal Reserve in 2005: "a lender does not have to report HMDA data unless it has an office in a metropolitan statistical area (MSA). As a result, reporting of home loans in some rural areas may be relatively low."

the Federal Reserve as a deposit-taking institution.

More specifically, financial institutions report any mortgage loans to their regulatory agency, and they are given a unique ID number and agency code in the loan-level data. The financial institutions we focus on are regulated by the Office of the Comptroller of the Currency (OCC, agency code 1), the Federal Reserve System (FRS, agency code 2), and the Federal Deposit Insurance Corporation (FDIC, agency code 3); thrifts regulated by the Office of Thrift Supervision (OTS, agency code 4); and independent mortgage companies regulated by the Department of Housing and Urban Development (HUD, agency code 7). This only leaves out credit unions, who make up a small portion of lending, and institutions regulated by the Consumer Financial Protection Bureau, an agency that was only created at the end of the decade as a response to the financial crisis. What we call banks are those institutions regulated by agencies corresponding to agency codes 1-4 since these are associated with a depository institution. Non-banks are non-depository independent mortgage lending companies and correspond to agency code 7.

We link lenders in the HMDA data to regulatory data filed by banks in the Report of Condition and Income (the Call Report) and by thrifts in the OTS Thrift Financial Report (TFR). Financial institutions that submit one of these forms are given a unique ID. If a bank (thrift) is part of a multi-bank (multi-thrift) holding company then each form provides a holding company ID which corresponds to the regulatory high holding company of the institution. For our analysis, we use the bank-holding or thift-holding company ID when an institution is part of a holding company and its unique ID otherwise. The HMDA ID is used for independent mortgage companies. Using a key of the HMDA Report IDs and the Call Report and TFR bank IDs provided to us by the Federal Reserve Board, we match the loan-level HMDA data to the bank level report data.[8]

The lender-level data set is the basis of the MSA and zip-code level data sets. Given the reporting restriction for originators in the HMDA data, we isolate our sample to zip codes

[8]We are grateful to Neil Bhutta who provided us access to this key.

that are located within MSAs. As of 2002, zip codes within MSAs account for 94% of the mortgage originations in HMDA. For these zip codes, we aggregate all HMDA originations by year, which gives us a zip-year level data set on mortgage originations. We also calculate for each zip code the 2002 non-core liability lender share (the NCL ratio). The 2002 NCL ratio is the weighted average NCL ratio of all lenders originating mortgages in the zip code in 2002, where the weights are determined by the total amount of originated mortgages by a given lender. We construct an MSA-level data set using the same procedure.

The analysis also uses individual-level credit bureau data from TransUnion, available through the Kilts Center at Chicago Booth. The TransUnion data are a 10% random sample of the universe. The analysis here uses the 2000 to 2007 files. From this data set, we construct at the zip code-year level the number of individuals with 1 or more mortgages, 3 or more mortgages, and 5 or more mortgages. We also use total mortgage debt outstanding at the zip-year level and the share of individuals with a Vantage Score below 660.

The other zip-code and MSA-level data sets are standard in the literature. The data sets include CoreLogic house price data at the zip code and MSA level. New units constructed come from the Census Building Permits Survey, which are available only at the county-level. As a result, we do not have a measure of construction at the zip code level. Total volume comes from the DataQuick data base, subsequently purchased by CoreLogic. The data sets we use also include a measure of cash-out refinancing volume as reported from mortgage refinancing data from CoreLogic. The CoreLogic refinancing data is described in more detail in Mian and Sufi (2014).

2.2 Summary statistics

Table 1 provides summary statistics for the lender level, zip code level, and MSA level data sets. The average ratio of non-core deposit liabilities to total liabilities is 0.74. Recall that this is defined to be one for non-bank mortgage lenders. Non-bank mortgage lenders make up 25% of the lender-level sample.

At the zip-code level, the 2002 NCL share is on average 0.77. The variables Δy_{BOOM} are constructed as follows. First, we add outcome y for zip code z in MSA m for years 2004 through 2006, and we then add outcome y for years 2000 through 2002. Δy_{BOOM} is defined to be the log difference between the two. Housing supply elasticity comes from Saiz (2010) and is available at the MSA level. For the number of new housing permits, we measure total construction during the boom as total units constructed in an MSA from 2004 to 2006, scaled by total housing units in the MSA as of 2000.

3 Empirical Strategy

The empirical strategy exploits a cross-sectional source of variation and a time-series source of variation. More specifically, it uses variation across geographic areas in exposure to mortgage lenders that typically rely heavily on non-core deposits in their liability structure. It then interacts this variation with acceleration in the PLS market in 2003 in the United States. This section explains the empirical strategy in more detail and explores both sources of variation.

3.1 Acceleration of the PLS market

The acceleration of the PLS market from 2003 to 2006 (shown in Figure 1) has been the topic of a large body of research (e.g., Chernenko et al. (2014)), and the consensus in this literature is that it reflected a supply-side phenomenon (e.g., Levitin and Wachter (2013); Justiniano et al. (2017)). The rapid rise of the PLS market was associated with a large increase in the quantity of mortgage originations and a sharp drop in mortgage interest spreads, indicative of an outward shift in mortgage credit supply (e.g., Justiniano et al. (2017); Demyanyk and Van Hemert (2011)).

The rise of the PLS mortgage market was part of the broader global pattern of the rise of securitization and shadow banking during the late 1990s and 2000s (e.g., Gorton and Metrick

(2012); Gorton and Metrick (2013)). In fact, there is evidence that the rise of asset-backed securitization during this time period was broader than the housing market. In Appendix Figure 1, we show that originated amounts in non-mortgage-related asset-backed securities increased from less than $300 billion to almost $800 billion from 2002 to 2006. It is unlikely that specific views on the U.S. housing market can explain the rise in these alternative markets. The empirical strategy does not take a stand on the precise source of the aggregate credit supply shock driving the PLS market during this time-frame. Researchers have put forth a number of explanations including a global savings glut (Bernanke (2005)), a rise in income inequality, neglected risks by investors (Gennaioli et al. (2012)), or lower uncertainty.

The specific timing of when the PLS market accelerated is discussed in Justiniano et al. (2017). They show that the Federal Reserve signaled higher interest rates in the summer of 2003, which led to a collapse in mortgage refinancing for conforming GSE-backed mortgages. Justiniano et al. (2017) argue that mortgage originations for the PLS market accelerated directly after this episode, and mortgage interest spreads over Treasuries fell sharply. As shown below, this is exactly the same time period in which lenders that rely on non-core deposits in their liability structure begin expanded credit more than other lenders.

3.2 Lender-level exposure to the PLS market

The second source of variation we use is lender-level exposure to the PLS market. Financial institutions rely on a number of sources of financing when originating loans. Research suggests that there is a critical distinction between institutions that rely on core deposits versus non-core liabilities (e.g., Hanson et al. (2015)). In particular, financial institutions that rely heavily on core deposits have a liability structure that is less prone to runs and cost shocks due to monetary policy (e.g., Hanson et al. (2015); Drechsler et al. (2017)). In return, they must hold costly equity capital and tend to invest in more illiquid assets. Further, the use of core deposits is closely related to an institution's ability to attract deposits from local customers where branches are available (e.g., Becker (2007)).

The idea behind our cross-sectional approach is that the acceleration of the PLS market from 2003 to 2006 represented a relative decline in the cost of funds for financial institutions that traditionally relied on liabilities other than core deposits. As a result, financial institutions that relied more heavily on non-core deposit financing in their liability structure as of 2002 experienced a relative increase in mortgage lending growth from 2002 to 2006 fueled by the ability to place mortgages into the PLS market. We call these institutions "high NCL lenders."

There are two sub-groups that make up the group of high NCL lenders: traditional banks with a high fraction of non-core deposits in their liability structure ("high NCL banks") and non-bank mortgage lenders ("non-banks"). For banks, we define the NCL ratio as of 2002 as one minus the ratio of core deposits to total liabilities, where core deposits are defined to be FDIC-insured deposits. Non-bank mortgage lenders rely completely on non-core deposit liability financing, and we therefore assume an NCL ratio of 1 for this group. In terms of specific financial institutions, notable high NCL banks as of 2002 were Countrywide Bank NA, JPMorgan Chase BK NA, and IndyMac BK FSB. Notable non-bank mortgage lenders as of 2002 were Ameriquest Mortgage Company, New Century Mortgage Corp, and American Home Mortgage Company.

3.3 Growth in mortgage originations: lender-level specifications

High NCL banks as of 2002 were more likely sell mortgages into the PLS market from 2002 to 2005. Figure 2 limits the sample to traditional banks (i.e., non-bank mortgage lenders are excluded) and plots the fraction of mortgages originated that are sold to a private institution. The outcome variable comes from the HMDA, which requires lenders to report to whom an originated loan is sold if it is sold within one year of origination. We follow Mian and Sufi (2009) and group together five categories that are a rough measure of mortgages sold into the PLS market.[9]

[9]These categories are mortgages sold (1) into private securitization, (2) to a commercial bank, savings bank, or savings affiliation affiliate, (3) to a life insurance company, credit union, mortgage bank, or finance

We split banks into two groups based on the NCL ratio as of 2002. The groups are weighted by total mortgage amount originated as of 2002, and the mean fraction sold is also weighted by total mortgage amount originated as of 2002. As Figure 2 shows, banks with a high NCL ratio of 2002 see a rapid rise in the fraction of originated mortgage amount sold to a private institution. The ratio increases from 20% in 2002 to almost 50% as of 2005. For banks with a low NCL ratio as of 2002, there is almost no change in the fraction of originated mortgage amount sold to a private institution.

Column 1 of Table 2 presents a lender-level regression of the change in the fraction of originated mortgage amount sold to a private institution from 2002 to 2005 on the NCL ratio as of 2002. The NCL ratio as of 2002 is divided by the sample standard deviation for ease of interpretation. This regression is limited to traditional banks; non-bank mortgage lenders are excluded. The coefficient estimate implies that a one standard deviation increase in the 2002 NCL ratio leads to a 15 percentage point increase in the share of originated mortgage amount sold to a private institution from 2002 to 2005.

In columns 2 through 4, we present or main lender-level regression specifications relating growth in originated mortgage amount to the 2002 NCL ratio. In column 2, we include non-bank mortgage lenders. By definition, the NCL ratio of a non-bank mortgage lender is 1. As the coefficient estimate in column 2 shows, a one standard deviation increase in the 2002 NCL ratio is associated with originated mortgage amount growth from 2002 to 2005 that is 18% higher.

The regressions reported in columns 3 and 4 explore the importance of non-bank mortgage lenders in explaining this correlation. Column 3 reports a regression specification in which we include an indicator variable for a non-bank mortgage lender. On average, non-bank mortgage lenders experienced an increase in mortgage lending that is 28% higher than banks from 2002 to 2005. In column 4 we include both the indicator variable and the 2002 NCL ratio. As it shows, the 2002 NCL ratio predicts originated mortgage amount growth even

company, (4) to an affiliate institution, or (5) to other type of purchaser.

with the inclusion of a non-bank mortgage lender indicator variable. Recall that the 2002 NCL ratio is 1 for all non-bank mortgage lenders; therefore, the statistically insignificant and small coefficient on the non-bank lender indicator implies that growth in originated mortgage amount for non-bank mortgage lenders is not statistically different from the linear prediction based on the 2002 NCL ratio of 1. The results in column 5 and 6 show that there is no pre-trend: high NCL lenders begin expanding amounts originated concurrent with the acceleration of the PLS market.

Figure 3 presents evidence on originated mortgage amount growth from 2000 to 2010 based on the 2002 NCL ratio. The top two panels represent the average amount originated by high and low NCL lenders, where the two groups represent lenders above and below the median 2002 NCL ratio. As before, both the averages and the groups are formed using the 2002 total amount originated as weights. The top left panel examines total amount originated whereas the top right is limited to mortgages for new home purchase. Both show a similar pattern. There is almost no difference between the two groups through 2002. Starting in 2003 and accelerating rapidly in 2004 and 2005, high NCL lenders see a relative expansion in mortgage originations.

In the bottom left panel, we present coefficient estimates $\{\beta_k\}$ from the following regression specification:

$$ln(y_{b,t}) = \alpha_b + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{b,2002} + \varepsilon_{b,t} \tag{1}$$

The left hand side variable is the natural logarithm of total amount originated by lender b in year t. The coefficient estimates $\{\beta_k\}$ provide the relative growth in mortgage amount originated by high NCL lenders. As the coefficients show, there is no pre-trend and a sharp relative rise for high NCL lenders starting in 2003 and accelerating during 2004 and 2005.

The evidence also suggests that there is a relative decline in mortgage lending by high NCL lenders when the PLS market collapses in 2007. However, this decline is underestimated because high NCL lenders are more likely to disappear from the sample after 2006. If a lender disappears, then it is not included in the sample for that year in the bottom left panel. The bottom right panel presents regression coefficients for a linear probability model that is similar to equation 1 except the left hand side variable is the probability of the lender being absent from the HMDA data in that year. As it shows, a one standard deviation increase in the 2002 NCL share implies a 10% higher probability of disappearing from the sample in 2007.[10]

3.4 Geographic exposure and MSA, zip code fixed effects

The rise of the PLS market led high NCL lenders to increase mortgage originations significantly more than low NCL lenders starting in 2003. The empirical strategy uses geographic variation across zip codes and MSAs in exposure to the high NCL lenders. For each geographic area in our sample, we calculate the 2002 NCL ratio as the average of the NCL ratios of mortgage lenders in the area, where the average is weighted by the amount of mortgage originations in 2002.

It is important to emphasize from the outset that these are not random areas of the country. Table 3 presents univariate regression coefficients for a set of observable variables regressed on the NCL ratio of an area in 2002. The first column shows the MSA-level coefficients, and the second column shows the zip code-level coefficients, where we include MSA fixed effects. High NCL ratio lenders have a higher market share of mortgage originations in MSAs with a lower deposit to mortgage origination ratio. We use this as a measure of the degree to which the MSA is "deposit-poor," and therefore must rely on funding from outside the MSA.

These MSAs also tend to have less elastic housing supply with higher average house prices. There is no significant correlation between the NCL share and income at the MSA

[10] In Appendix Figure 2, we present results separately for refinancing originations. The results are similar: there is no significant pre-trend, and high NCL share lenders see stronger relative growth in refinancing originations starting in 2003.

level. However, within MSAs, high NCL lenders have the largest market share in lower income zip codes. The sign flips for home values. While MSAs with a high NCL exposure have higher average house prices, the high NCL lenders appear to focus within these MSAs on zip codes with lower house prices.

For the rest of the relationships, the across MSA and within MSA coefficients have the same sign. Areas with high NCL exposure tend to have lower homeownership rates, lower credit scores, and a younger population. The fact that high NCL lenders have large market share in zip codes with a lower fraction of individuals over the age of 65 is consistent with Becker (2007), who shows that seniors tend to save via deposits in local banks. Older zip codes are therefore "deposit-rich," and are less reliant on outside sources of funding. All of these correlations make economic sense: deposit-poor areas are more likely to rely on mortgages originated by lenders that rely on external funding. In this sense, we do not want to control for these factors; they are the underlying source of variation in exposure to high NCL lenders.

However, this raises a concern: are shocks in these areas other than the acceleration of the PLS market responsible for the patterns we find? We examine this concern by focusing on lender-MSA and lender-zip code level data sets. This allows us to include MSA or zip code fixed effects when estimating the effect of a high 2002 NCL share on amount originated growth at the lender level. More specifically, Table 4 presents estimates from the following equation:

$$\Delta y_{b,g,0205} = \alpha_g + \beta NCL_{b,2002} + \varepsilon_{b,g,0205}$$

where the outcome variable is the growth in originated mortgage amount by lender b in geography g from 2002 to 2005. The geographical unit is an MSA in columns 1 and 2 of Table 4 and a zip code in columns 3 and 4.

Column 1 reports the MSA-lender level specification without fixed effects, which is similar to the estimate reported in column 3 of Table 2. In column 2, we report the specification with MSA fixed effects. The R^2 increases from 0.04 to 0.16, which indicates the statistical power of the MSA fixed effects in capturing variation in lender originations. However, the point estimate on the 2002 NCL share drops only slightly.

Columns 3 and 4 conduct the same estimation at the zip code-lender level. The inclusion of zip code fixed effects boosts the R^2 by a factor of 6. The point estimate declines slightly, but it remains economically large and statistically significant at the 0.001 level. If anything, this specification may "over-control" for the effect of the acceleration of the PLS market. If there is any spillover onto other lenders lending in the same zip code from the acceleration of the PLS market, the specification reported in column 4 will eliminate this effect. We view the estimate in column 4 as a particularly convincing result that the expansion of lending by high NCL lenders is not due to differential geographical exposure in their lending markets. The relative expansion of mortgage lending by high NCL share lenders appears to be due to the acceleration of the PLS market as opposed to local economic conditions.

4 Mortgage Originations, Volume, and Investors

Areas of the country with greater exposure to high NCL share lenders as of 2002 witnessed stronger growth in originated mortgage amounts, housing transaction volume, and the number of individuals with many mortgages from 2002 to 2005. In all of the subsequent analysis, the main right hand side variable is the average NCL share of mortgage lenders in the area as of 2002, where the average is weighted by the total amount of mortgage originations by the lender in the area in 2002.

4.1 Mortgage originations

Figure 4 splits zip codes in the sample into high and low NCL share zip codes based on the population-weighted median, and it plots home purchase mortgage origination amount and refinancing origination amount indexed to 2002. As it shows, there is no pre-trend from 2000 to 2002. From 2003 to 2005, high NCL exposure zip codes see stronger growth in amounts originated for home purchase and for refinancing. The relative expansion reverses from 2006 to 2010.

In Figure 5, we present coefficients from the following specification:

$$ln(y_{z,t}) = \alpha_z + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{z,2002} + \varepsilon_{z,t} \tag{2}$$

The left hand side variables are the natural logarithm of the same two mortgage measures in Figure 4 in zip code z in year t. The coefficients β_k trace the relative growth of originated amounts in zip codes with a high NCL share as of 2002. As in all specifications, $NCL_{z,2002}$ is normalized to have a standard deviation of one to ease interpretation.

Zip codes with high NCL exposure as of 2002 witness strong relative growth in both measures of originated mortgage amounts from 2002 to 2006. Mortgage originations subsequently collapse after 2006, and by 2010 they are far below the 2002 level. Our interpretation of this pattern is that the PLS market created a more amplified mortgage origination cycle from 2002 to 2010 in high NCL share zip codes.

Figure 6 presents coefficients for a specification similar to equation 2, except it uses MSA by month level data and the outcome variable is total mortgage amounts originated. The higher frequency of the data shows the precise timing of the relative rise in mortgage origination growth in high NCL share MSAs. The coefficient rises sharply in September and October of 2003.

The right panel zooms in on 2003, and it also includes the PLS mortgage spread to Treasury rate residual from Justiniano et al. (2017). They show a sharp decline in the PLS spread in August through October of 2003, which they associate with an aggressive expansion in mortgage originations by PLS lenders. The relative rise in amount originated in high NCL share MSAs starts at almost the exact same time as the PLS spread drops. Our interpretation of this pattern is that the acceleration of the PLS market lowered mortgage interest spreads and led to a sudden relative rise in originations in high NCL share MSAs. The high frequency analysis supports the view that high NCL share MSAs experienced a sudden rise in originations because of the acceleration of the PLS market as opposed to a change in income or beliefs.

Table 5 presents regression coefficients from the following specification:

$$\Delta y_{z,m,BOOM} = \alpha_m + \beta NCL_{z,m,2002} + \varepsilon_{z,m} \tag{3}$$

The outcome variable $\Delta y_{z,m,BOOM}$ is constructed as follows. First, we add outcome y for zip code z in MSA m for years 2004 through 2006, and we then add outcome y for years 2000 through 2002. $\Delta y_{z,m,BOOM}$ is defined to be the log difference between the two. In other words, $\Delta y_{z,m,BOOM}$ is the log difference in the three year sum of the outcome from the boom period less the pre-boom period, where we exclude 2003 as a transition year. This specification is meant to capture the differential cumulative flow of originated mortgage amount during the boom period relative to the pre-boom period.

The coefficient estimates in Table 5 imply a positive and statistically significant effect of the 2002 NCL share in a zip code on the subsequent increase in mortgage amount originated. In the even columns, we add MSA fixed effects. The coefficient estimates on the 2002 NCL share variable become larger with the addition of MSA fixed effects. The MSA fixed effects are boosting the R^2 by at least a factor of four across the specifications, and their inclusion is actually boosting the estimated coefficient on the 2002 NCL share. This suggests that MSA-level omitted shocks are not responsible for the effect of the NCL share on originated mortgage amounts. In terms of magnitudes, the MSA fixed effects specification imply that a one standard deviation increase in the 2002 NCL share leads to a 12% increase in amounts originated for home purchase and a 29% increase refinancing amount originated.[11]

[11] Appendix Figure 3 shows that high NCL share zip codes experience a large relative increase in the share of mortgages refinanced with cash taken out.

The specifications reported in columns 5 and 6 use data from BlackBox Logic on the number of mortgages originated in a zip code that were subsequently placed into the PLS market. As the specifications show, high NCL share zip codes see significantly stronger growth in origination of mortgages subsequently placed into the PLS market. This is consistent with the view that it is the expansion of the PLS market that leads to higher mortgage origination growth in high NCL share zip codes.

Columns 7 and 8 report specifications using the growth non-PLS mortgage originations, where non-PLS mortgage originations are total HMDA originations minus BlackBox originations. High NCL share zip codes see stronger growth in non-PLS mortgage originations, but the coefficients are 40 to 60% smaller than the growth in PLS mortgage originations. There are a number of factors that explain why even non-PLS mortgage origination growth is stronger in high NCL share zip codes. According to Piskorski et al. (2015), the BlackBox Logic data set captures only 90% of the mortgages sold into the PLS market. Many of the mortgages we label as non-PLS may actually be PLS mortgages. In addition, BlackBox contains mortgages successfully placed into the PLS market; high NCL lenders may originate mortgages for the purpose of PLS which subsequently are not placed into the market. Finally, the higher house price growth induced by the expansion of mortgage originations by high NCL lenders leads to a spike in refinancing activity in high NCL share zip codes as shown in the appendix. Some of this refinancing activity likely took place in the non-PLS market.

4.2 Volume

Figure 7 presents coefficients from the estimation of equation 2 using log number of home purchase mortgage originations and log number of transactions as the left hand side variables. Both panels show a large relative increase in the volume of housing transactions from 2003 to 2006 in high NCL share zip codes. The relative increase begins exactly as the PLS market accelerates in 2003. A one standard deviation increase in the 2002 NCL share of a zip code leads to a 10% increase in the number of purchase mortgages and a 20% increase in transaction volume.

Table 6 presents estimates of equation 3 using transactions as the left hand side variable. More specifically, the left hand side variable is the number of transactions in a zip code from 2004 to 2006 minus the number from 2000 to 2002, and the difference is scaled by the housing stock as of 2000. The specifications reported in Table 6 include MSA fixed effects; the specifications without MSA fixed effects are in Appendix Table 1.

As column 1 shows, a one standard deviation increase in the 2002 NCL share of a zip code leads to a 6 percentage point increase in transaction volume as a share of the housing stock. We split this effect into four sub-components based on the mailing address listed by the buyer in the transaction. This is a common technique to measure whether the house is purchased by an occupier or an investor. If the mailing address listed on the transaction is different than the address of the home, then the assumption is that the buyer does not intend to live in the house.

Just over 50% of the total effect is from buyers that report a mailing address that is the same as the purchased home. There is a substantial relative increase in volume of transactions with the mailing address missing. Based on these findings, it is difficult to know whether the houses being purchased are purchased for investment or for occupancy. As shown in Piskorski et al. (2015), Griffin and Maturana (2016b), and Griffin and Maturana (2016a), the PLS market was plagued with fraud in which buyers were listed as owner-occupiers even though they were investors. It seems likely that such fraud encouraged mortgage originators and buyers to report the same mailing address on the deed transfer even if the house was purchased as an investment.

4.3 Investors

Figure 8 explores an alternative measure of investment purchases in a zip code. More specifically, it plots coefficients for specifications identical to equation 2 with the left hand side variable being the log number of individuals with 1 or more, 3 or more, or 5 or more mortgages. The assumption underlying the analysis is that an individual is unlikely to have 5 or more mortgages on the same property, and so the individual is likely buying multiple properties.

The top left panel shows that high NCL share zip codes do not see a disproportionate rise in the number of individuals with one or more mortgages. Instead, the top right and bottom left panels show a sharp relative rise in the number of individuals with 3 or more mortgages, or 5 or more mortgages. A one standard deviation increase in the 2002 NCL share leads to 5% higher growth in the numbe of individuals with 5 or more mortgages.

Columns 1 through 6 of Table 7 show the corresponding regression table. There is no significant relative increase in the number of individuals with one or more mortgages in high NCL share zip codes, but there is a large relative increase in the number of individuals with 3 or more, or 5 or more mortgages. Inclusion of MSA fixed effects leads to a substantial increase in the coefficient estimates. The coefficient estimate in column 6 implies that a one standard deviation increase in the 2002 NCL share leads to a 9% increase in the number of individuals with 5 or more mortgages, which is 1/3 a standard deviation of the left hand side variable.

Figure 8 and Table 7 also explore the effect of the acceleration of the PLS market on total mortgage debt outstanding. There is a positive effect, but the size is modest and it is not robust to the inclusion of MSA fixed effects. This shows that the acceleration of the PLS market is not directly responsible for the rise in household debt in the United States from 2000 to 2007.

5 House Prices, Construction, and Boom MSAs

5.1 House prices

The acceleration of the PLS market at the end of 2003 led to a significant relative boost in mortgage originations and volume in high NCL share zip codes. What was the effect on house prices and construction activity? The left panel of Figure 9 presents estimates of β_k from the estimation of equation 2 with the logarithm of house prices as the left hand side variable. As the estimates show, high NCL share zip codes experience positive relative growth in house prices in 2003, which then accelerates rapidly in 2004, 2005, and 2006. The estimates imply that a one standard deviation in the 2002 NCL share of a zip code leads to an 8 percentage point increase in house prices from 2002 to 2006.

The PLS market collapsed in 2007, which corresponds to a collapse in house prices in high NCL share zip codes. In fact, the collapse is severe enough that the log house price level ends up lower in 2009 and 2010 than its 2002 level relative to low NCL share zip codes. Recall that this is a relative coefficient estimate, and so the negative coefficient estimates in 2009 and 2010 imply that the log house price level difference in high NCL share zip codes relative to low NCL share zip codes is lower in 2010 than it was 2002. We will return to this point below.

Table 8 presents regression estimates from the following equation:

$$\Delta HP_{z,m,2002,2006} = \alpha_m + \beta NCL_{z,m,2002} + \gamma SAIZ_m + \delta NCL_{z,m,2002} SAIZ_m + \varepsilon_{z,m} \qquad (4)$$

We include housing supply elasticity as a control variable in all specifications because of the standard relationship we would expect between house price growth and a demand shock caused by increased credit availability: for the same shock in demand, one would expect house price growth to be stronger in MSAs with more inelastic housing supply (Glaeser et al. (2008)). We follow the literature and use the Saiz (2010) measure of housing supply elasticity at the MSA level. Columns 1 through 3 present estimates of the specification without MSA fixed effects, and columns 4 and 5 present estimates without MSA fixed effects.

As the estimate in column 2 shows, the 2002 NCL share in a zip code predicts high house price growth from 2002 to 2006. In terms of magnitudes, a one standard deviation increase in the 2002 NCL share leads to 6% higher house price growth. Column 4 includes the interaction between the 2002 NCL share and the measure of housing supply elasticity. As expected, the coefficient is negative: the effect of 2002 NCL share on house price growth is weaker for zip codes located in more elastic housing supply MSAs.

The estimate on the 2002 NCL share variable in column 4 shows that the effect of the 2002 NCL share on house price growth is significantly stronger in the most inelastic housing supply MSAs. A one standard deviation increase in the 2002 NCL share boosts house price growth by 15% from 2002 to 2006 in the most inelastic housing supply MSAs. In MSAs where it is difficult to expand the supply of housing, the demand shock induced by the acceleration of the PLS market pushes house prices higher.

In columns 5 and 6, we include MSA fixed effects in order to examine only within-MSA zip code level variation in house prices.[12] The results are qualitatively similar: high 2002 NCL share zip codes see stronger growth in house prices relative to low 2002 NCL share zip codes located within the same county. Further, this effect is weaker in more elastic housing supply MSAs.

However, the absolute value of the coefficients drop between one-third to one-half in the specifications with MSA fixed effects. This is likely due to both statistical and economic reasons. The statistical reason can be seen with an examination of the increase in the R^2 when MSA fixed effects are included. For example, comparing columns 2 and 5, the R^2 increases from 0.132 to 0.930. Such a large increase in the R^2 suggests that CoreLogic is smoothing its zip code level house price indices within MSAs.[13] Such smoothing in the

[12]The Saiz (2010) elasticity measure is defined only at the MSA level, and so the elasticity level drops out of the MSA fixed effects specification.

[13]We are currently examining the documentation to explore this issue further.

outcome variable within MSAs would reduce the coefficient estimate on the 2002 NCL share when including MSA fixed effects, even if in reality the coefficient estimate would be the same in the absence of smoothing. The economic reason is the presence of spillovers. Zip codes within an MSA are not isolated islands. Price effects will therefore be muted as potential buyers search across neighboring zip codes. Such spillovers suggest that the use of within-MSA variation may lead to an underestimate of the effect of the NCL share on house price growth.

5.2 Construction

The right panel of Figure 9 examines construction activity as measured by new building permits collected by the Census. Data on new building permits are available only at the county level, and so we estimate equation 2 at the MSA level instead of the zip code level.

As the estimates show, there is a relative rise in construction activity in high NCL share MSAs that starts in 2003 and accelerates in 2004 through 2006. In terms of magnitudes, a one standard deviation increase in the 2002 NCL share of an MSA leads to an 7 to 8 percentage point increase in units constructed from 2002 to 2006. As with house prices, construction activity begins falling in 2007 and then collapses by 2009 and 2010. Relative to low NCL share MSAs, construction activity falls below the 2002 level in high NCL share MSAs.

Table 9 presents estimates from a regression specification similar to equation 4 above with two differences. First, the specifications are estimated at the MSA level instead of the zip code level. Second, the left hand side variable is cumulative number of housing units constructed from 2004 to 2006 scaled by the number of housing units in the MSA in 2000. The estimate reported in column 2 implies that a one standard deviation increase in the 2002 NCL share of an MSA is associated with a 1 percentage point increase in the number of units constructed from 2004 to 2006 as a fraction of the 2000 housing stock. This is a 0.4 standard deviation of the left hand side variable. As column 3 shows, the coefficient estimate is almost identical when including housing supply elasticity of the MSA as a control variable.

In column 4, we interact the 2002 NCL share with housing supply elasticity. The coefficient estimate on the interaction is statistically different than zero. This is a puzzle. We would expect a stronger housing demand shock induced by the acceleration of the PLS market to induce more construction in MSAs with elastic housing supply. We believe this result is related to findings in Glaeser et al. (2008), Davidoff (2013), and Nathanson and Zwick (2017) who find evidence of anomalous MSAs that experienced both high house price growth and construction growth from 2002 to 2006. We explore this issue in more detail in the next sub-section.

5.3 The anomalous bubbly MSAs

Motivated by Glaeser et al. (2008), Davidoff (2013), and Nathanson and Zwick (2017), we define a bubble MSA as one that experiences a simultaneous large rise in construction activity and house prices from 2002 to 2006. The left panel of Figure 10 plots house price growth from 2002 to 2006 against construction activity during the boom for the MSAs in the sample. Both measures are standardized to be mean zero and standard deviation one. The 45 degree line is also plotted. A given MSA's bubble measure is constructed by drawing a perpendicular line from the MSA dot to the 45 degree line, and then measuring the distance from that intersection to the (0,0) point on the graph.[14] The higher the bubble measure, the higher the simultaneous increase in house prices and construction in an MSA. The right panel shows the top 20 MSAs in terms of the bubble measure. These are relatively elastic housing supply cities that witnessed a simultaneous large increase in house price growth and construction activity during the boom.

Glaeser et al. (2008) build a model to understand the existence of MSAs with a simultaneous increase in house prices and construction. As they put it, "rational bubbles can exist when the supply of housing is fixed, but not with elastic supply and a finite number

[14]More formally, all points on the graph in the left panel of Figure 10 are rotated 315 degrees using standard trigonometry formulas, and the vertical distance from zero is the bubble measure.

of potential home buyers." They examine a model with an irrational temporary increase in optimism about future prices, and they show that such a shock can help explain why prices and construction may increase rapidly in MSAs with elastic housing supply. They also make a prediction that the crash will be particularly severe in these MSAs, because there will be overbuilding during the bubble.

Our contribution to this discussion is to show that MSAs with a high NCL share as of 2002 are the most likely to have both strong house price and construction growth during the subsequent housing boom. That is, what happened in bubbly MSAs seems directly related to the acceleration of the PLS market from 2003 to 2006.

Figure 11 plots the bubble measure against the NCL share of the MSA as of 2002. There is a strong positive relationship. The positive relationship is confirmed in the first column of Table 10. A one standard deviation increase in the NCL share of an MSA leads to a 0.4 increase in the bubble measure, which is one-third a standard deviation. The specification in column 2 includes fixed effects for the nine Census divisions given that the bubbly MSAs tend to be concentrated in the West and Southeast regions of the country. The coefficient estimate falls by one-half, but it remains statistically significant and large.

Columns 3 and 4 use a slightly different specification, where the left hand side variable is equal to one if an MSA ends up in the top quartile of both the house price growth and construction growth distribution and zero otherwise. As column 3 shows, a one standard deviation increase in the NCL share as of 2002 leads to an 8 percentage point increase in the probability of ending up in the top quartile of both distributions. This is a large effect given that the mean of the left hand side variable is 10%.

Table 10 also tests an additional prediction from the Glaeser et al. (2008) model: the collapse in house prices in boom MSAs may bring house prices even below their initial level before the boom because of overbuilding. To test this prediction, columns 5 through 8 report specifications relating the long run change in house prices and construction to the NCL share as of 2002. The underlying assumption is that the housing market was in a steady state equilibrium in 2002 before the acceleration of the PLS market. If housing supply were completely inelastic, an unexpected demand shock and reversal would lead to a boom and bust in house prices, but the long run level would remain constant at the 2002 level.

As columns 5 through 8 show, house price and construction growth from 2002 to 2010 were relatively lower in MSAs with a high 2002 NCL share. These results are robust to the inclusion of Census division fixed effects. The estimates suggest that the acceleration of the PLS market from 2002 to 2006 led to housing market excesses in high NCL share MSAs that eventually caused house prices to fall even below their pre-boom house price level.

6 Housing Market Optimism

One explanation of the housing boom put forth in the literature is a change in expectations of house price growth. How is housing market optimism related to the acceleration of the PLS market in 2003? This section explores this question.

6.1 Data and aggregate patterns

Large sample data sets on beliefs about the evolution of house prices and the housing market prior to 2007 are unavailable.[15] Here, we follow Piazzesi and Schneider (2009) and use the Michigan Survey of Consumers to measure beliefs about the housing market. In particular, the Michigan survey asks the following question to respondents: "Generally speaking, do you think now is a good time or bad time to buy a house?" Almost 99% of survey respondents answer this question by saying it is either a "good" or "bad" time. In addition to this question, respondents are asked a follow up question: "Why do you say so?" In response to this question, respondents give up to two answers. The survey questioners record a number of different responses, which we classify into sub-groups as follows.

If an individual responds that now is a good time to buy a house because "prices are

[15] To the best of our knowledge, the only data set that records house price expectations prior to 2007 in the United States is used in Case et al. (2012) and covers only four metropolitan areas.

going up," "prices are low/stable/not too high," "prices won't get any lower," or "good buys available," we classify the response as being favorable because of house price considerations. If an individual answers that now is a bad time to buy a house because "prices will fall later, will come down" or "prices are too high, houses cost too much," we classify the response as being unfavorable because of house price considerations. Note that there are many other responses to the follow up question; the most common response given is that now is a good time to buy a home because "interest rates are low." As a result, the fractions of respondents saying now is a good time versus a bad time to buy because of house prices considerations will not add up to one. Further, the total fraction mentioning prices as a consideration will vary over time.

Figure 12 shows that evolution of answers to the Michigan questions from 2000 to 2010 for the full sample. During the PLS acceleration period of 2003 to 2006, the fraction of individuals saying that now is a bad time to buy a house rises from 20% to almost 40%. This result matches the finding in Piazzesi and Schneider (2009). Further, this increase is driven in large part by individuals who are sour on the housing market because of high prices. These findings match those of Piazzesi and Schneider (2009), and they are difficult to reconcile with the view that general optimism about housing was responsible for the rise in house prices from 2003 to 2006.

However, consistent with Piazzesi and Schneider (2009), there is a cluster of individuals who appear to become more optimistic about housing because of price considerations. Although the total fraction of individuals saying it is a favorable time to buy a home falls, the fraction saying it is a good time to buy because of price consideration increases. Piazzesi and Schneider (2009) conclude based on this evidence that a small group of optimistic individuals can have a large effect on house prices.

6.2 MSA-level analysis

How do the aggregate patterns on housing market optimism relate to the acceleration of the PLS market? We focus on cross-sectional variation across MSAs in the 2002 NCL share to explore this question. To the best of our knowledge, this study is the first to examine cross-sectional variation across MSAs in the answers to home buying questions in the Michigan Survey of Consumers.[16]

Table 11 reports specifications on the evolution of optimism on the housing market in high versus low NCL share MSAs both before and during the housing boom. Columns 1 and 2 show that from 2000 to 2002, there is no statistically significant relative change in the fraction of individuals in an MSA with a high NCL share saying that it is a good time to buy a home because of price considerations. The standard errors are small; there is no evidence that a local housing market optimism shock precedes the rise of mortgage originatons, house prices, or construction in high NCL share MSAs.

To shed light on the relationship between house price growth and housing market optimism, columns 3 through 6 present estimates of the following specification:

$$\Delta MichMeasure_{m,BOOM} = \alpha_m + \beta * HPGrowth_{m,0206} + \varepsilon_{z,m}$$

where $\Delta MichMeasure_{m,BOOM}$ is the MSA-level average of the survey responses to a given Michigan question in MSA m in years 2004 through 2006 minus the average of the survey responses to the same question in MSA m in years 2000 to 2002. Columns 3 and 4 present the OLS estimates, and columns 5 and 6 present instrumental variable estimates where the instrument for house price growth is the 2002 NCL share of the MSA.

The OLS and IV estimates convey a consistent message: the average household in high house price growth areas becomes more pessimistic about the housing market in 2004 through 2006 relative to 2000 to 2002. In terms of magnitudes, a one standard deviation increase in

[16]The most disaggregated geographic identifiers available for the Michigan survey are at the county-level; zip code level analysis is therefore not possible.

house price growth leads to a 6 to 8 percentage point increase in the share of individuals expressing pessimism on the housing market. The increasing pessimism is driven by people who are pessimistic because of house price considerations. In Appendix Table 2, we split the "bad time to buy because of prices considerations" into the two separate subcomponent answers: "bad time to buy because prices are too high," and "bad time to buy because prices will fall." For both components, we find that there is a relative increase in the fraction of individuals expressing pessimism in high house price growth MSAs during the boom.

Overall, the results are inconsistent with the view that a general rise in housing market optimism in high NCL share MSAs instigated the housing boom in these cities. Instead, higher house price growth, fueled by the acceleration of the PLS market, made the average individual in these MSAs more pessimistic about house prices. This provides further evidence that the PLS market affected the housing market not through a general rise house price expectations, but instead through boosting the buying power of a smaller group of individuals.

7 Discussion

7.1 Relation to existing research

The findings presented here are related to a large body of research on the PLS market and subprime mortgages in particular (e.g., Keys et al. (2010); Mayer et al. (2009); Keys et al. (2012) Demyanyk and Van Hemert (2011); Purnanandam (2011); Dell'Ariccia et al. (2012); Piskorski et al. (2015); Griffin and Maturana (2016b); Griffin and Maturana (2016a)). The findings in this literature suggest that the PLS market was plagued with incentive problems, fraud, and poor underwriting, and represented a "classic lending boom-bust scenario" (Demyanyk and Van Hemert (2011)). To the best of our knowledge, this study is the first to isolate a plausibly exogenous source of cross-sectional variation in geographic exposure to the acceleration of the PLS market in 2003 in order to test how the rise of the PLS market affected house prices, construction activity, and housing market optimism.

There is a related body of research focusing on anomalous elastic housing supply MSAs with both a boom in construction and house prices (Glaeser et al. (2008); Davidoff (2013); Chinco and Mayer (2015); and Nathanson and Zwick (2017)). Nathanson and Zwick (2017) point to the importance of supply-side speculation and Chinco and Mayer (2015) point to the importance of out-of-town investors. The findings presented here are compatible with these channels, but they point to the acceleration of the PLS market as an instigating factor in explaining the anomalous bubbly MSAs such as Las Vegas and Phoenix.

This study is also related to the body of research exploring the role of speculation and investor purchases in the housing cycle of 2000 to 2010 (e.g., DeFusco et al. (2017); Chinco and Mayer (2015); Bhutta (2015); Haughwout et al. (2011)). We know for example that states that experienced the largest boom-bust cycle witnessed the largest increase in the participation of investors (Haughwout et al. (2011)), investors played an important role in explaining the rise in household debt levels (Bhutta (2015)), and short-term investors amplified volume and price movements in many markets (DeFusco et al. (2017)). The results presented here show that the acceleration of the PLS market was a key instigating factor in explaining the rise of speculation and investor purchases in the housing market, a hypothesis which has not been explored in the existing literature.

There is a body of research exploring the rise of household debt across the income and credit score distribution during the housing boom (Mian and Sufi (2009); Adelino et al. (2016); Mian and Sufi (2017a); Mian and Sufi (2017b); Adelino et al. (2017); Foote et al. (2016); Albanesi et al. (2017)). The focus of this research is on the rise in the level of household debt, which is not a primary outcome explored in this study. The findings presented above show that the acceleration of the PLS market does not directly explain the aggregate rise in household debt, which was driven by existing homeowners extracting equity during the housing boom (e.g., Mian and Sufi (2011); Mian and Sufi (2015); Mian and Sufi (2017b)).

The analysis here does not focus on the average income or average credit score of zip codes or MSAs; instead, the analysis defines the instrument as exposure to the PLS market, which is positively correlated with income across MSAs, but negatively correlated within MSAs. The results are consistent with some of the correlations in the existing literature. For example, high PLS share zip codes see substantial growth in volume and the number of home purchase mortgage originations from 2002 to 2005. Within MSAs, high PLS share zip codes tend to have lower income levels, and so the results are consistent with the findings of Adelino et al. (2016) and Foote et al. (2016) of higher mortgage origination "churn" rates in low income neighborhoods.

7.2 Future work

This study exploits cross-sectional variation in geographical areas across the United States in exposure to lenders that expand mortgage originations due to the acceleration of the PLS market in the summer of 2003. High PLS share areas witness strong relative growth in mortgage originations, volume, house prices, and construction from 2002 to 2006. High PLS share MSAs are more likely to become "bubbly" MSAs: cities that experience a simultaneous rise in both construction and house price growth. Areas more exposed to the acceleration of the PLS market experience significant relative growth in the number of individuals with many mortgages. Further, individuals living in these MSAs on average become more pessimistic on the housing market during the boom. These latter two findings suggest that credit supply expansion through the PLS market increased house price growth by increasing the buying power of a small cluster of individuals.

References

Adelino, Manuel, Antoinette Schoar, and Felipe Severino, "Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class," *Review of Financial Studies*, 2016, *29* (7), 1635–1670.

_ , _ , and _ , "Dynamics of Housing Debt in the Recent Boom and Great Recession," Technical Report, National Bureau of Economic Research 2017.

Albanesi, Stefania, Giacomo De Giorgi, and Jaromir Nosal, "Credit growth and the financial crisis: A new narrative," Technical Report, National Bureau of Economic Research 2017.

Becker, Bo, "Geographical segmentation of US capital markets," *Journal of Financial economics*, 2007, *85* (1), 151–178.

Bernanke, Ben, "The global saving glut and the US current account deficit," Technical Report, Board of Governors of the Federal Reserve System (US) 2005.

Bhutta, Neil, "The ins and outs of mortgage debt during the housing boom and bust," *Journal of Monetary Economics*, 2015, *76*, 284–98.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo, "Understanding booms and busts in housing markets," *Journal of Political Economy*, 2016, *124* (4), 1088–1147.

Case, Karl E, Robert J Shiller, and Anne K Thompson, "What Have They Been Thinking? Homebuyer Behavior in Hot and Cold Markets [with Comments and Discussion]," *Brookings Papers on Economic Activity*, 2012, pp. 265–315.

Chernenko, Sergey, Samuel G. Hanson, and Adi Sunderam, "The Rise and Fall of Demand for Securitizations," 2014. Working Paper.

Chinco, Alex and Christopher Mayer, "Misinformed speculators and mispricing in the housing market," *The Review of Financial Studies*, 2015, *29* (2), 486–522.

Davidoff, Thomas, "Supply Elasticity and the Housing Cycle of the 2000s," *Real Estate Economics*, 2013, *41* (4), 793–813.

DeFusco, Anthony A, Charles G Nathanson, and Eric Zwick, "Speculative dynamics of prices and volume," Technical Report, National Bureau of Economic Research 2017.

Dell'Ariccia, Giovanni, Luc Laeven, and Deniz Igan, "Credit Booms and Lending Standards: Evidence from the Subprime Mortgage Market," *The Journal of Money, Credit and Banking*, 2012, *44* (2-3), 367–384.

Demyanyk, Yuliya and Otto Van Hemert, "Understanding the subprime mortgage crisis," *Review of financial Studies*, 2011, *24* (6), 1848–1880.

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl, "The deposits channel of monetary policy," *The Quarterly Journal of Economics*, 2017, *132* (4), 1819–1876.

Foote, Christopher L, Lara Loewenstein, and Paul S Willen, "Cross-sectional patterns of mortgage debt during the housing boom: evidence and implications," Technical Report, National Bureau of Economic Research 2016.

Gao, Zhenyu, Michael Sockin, and Wei Xiong, "Economic consequences of housing speculation," Technical Report, Working paper 2016.

Geanakoplos, John, "The leverage cycle," *NBER macroeconomics annual*, 2010, *24* (1), 1–66.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny, "Neglected risks, financial innovation, and financial fragility," *Journal of Financial Economics*, 2012, *104* (3), 452–468.

Glaeser, Edward L, Joseph Gyourko, and Albert Saiz, "Housing supply and housing bubbles," *Journal of urban Economics*, 2008, *64* (2), 198–217.

Gorton, Gary and Andrew Metrick, "Securitized banking and the run on repo," *Journal of Financial economics*, 2012, *104* (3), 425–451.

_ **and** _ , "Securitization," in "Handbook of the Economics of Finance," Vol. 2, Elsevier, 2013, pp. 1–70.

Griffin, John M. and Gonzalo Maturana, "Did Dubious Mortgage Origination Practices Distort House Prices?," *The Review of Financial Studies*, 2016, *29*, 1671–1708.

_ **and** _ , "Who Facilitated Misreporting in Securitized Loans?," *The Review of Financial Studies*, 2016, *29*, 384–419.

Hanson, Samuel G, Andrei Shleifer, Jeremy C Stein, and Robert W Vishny, "Banks as patient fixed-income investors," *Journal of Financial Economics*, 2015, *117* (3), 449–469.

Haughwout, Andrew, Donghoon Lee, Joseph S Tracy, and Wilbert Van der Klaauw, "Real estate investors, the leverage cycle, and the housing market crisis," 2011.

Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti, "The Mortgage Rate Conundrum," 2017. Working Paper.

Keys, Benjamin J, Amit Seru, and Vikrant Vig, "Lender Screening and the Role of Securitization: Evidence from Prime and Subprime Mortgage Markets," *The Review of Financial Studies*, 2012, *25* (7), 2071–2108.

_ **, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig**, "Did securitization lead to lax screening? Evidence from subprime loans," *The Quarterly journal of economics*, 2010, *125* (1), 307–362.

Levitin, Adam and Susan Wachter, "Explaining the Housing Bubble," *Georgetown Law Journal*, 2013, *100* (4), 1177–1258.

Maggio, Marco Di and Amir Kermani, "Credit-induced boom and bust," *The Review of Financial Studies*, November 2017, *30* (11), 3711–58.

Mayer, Christopher, Karen Pence, and Shane Sherlund, "The Rise of Mortgage Defaults," *Journal of Economic Perspectives*, 2009, *23* (1), 27–50.

Mian, Atif and Amir Sufi, "The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis," *The Quarterly Journal of Economics*, 2009, *124* (4), pp. 1449–1496.

_ **and** _ , "House prices, home equity–based borrowing, and the us household leverage crisis," *The American Economic Review*, 2011, *101* (5), 2132–2156.

_ **and** _ , "House price gains and US household spending from 2002 to 2006," Technical Report, National Bureau of Economic Research 2014.

_ **and** _ , *House of debt: How they (and you) caused the Great Recession, and how we can prevent it from happening again*, University of Chicago Press, 2015.

_ **and** _ , "Fraudulent income overstatement on mortgage applications during the credit expansion of 2002 to 2005," *The Review of Financial Studies*, 2017, *30* (6), 1832–1864.

Mian, Atif R and Amir Sufi, "Household Debt and Defaults from 2000 to 2010: The Credit Supply View," *Evidence and Innovation in Housing Law and Policy*, 2017.

Nathanson, Charles G and Eric Zwick, "Arrested development: Theory and evidence of supply-side speculation in the housing market," Technical Report, National Bureau of Economic Research 2017.

Piazzesi, Monika and Martin Schneider, "Momentum traders in the housing market: survey evidence and a search model," *American Economic Review*, 2009, *99* (2), 406–11.

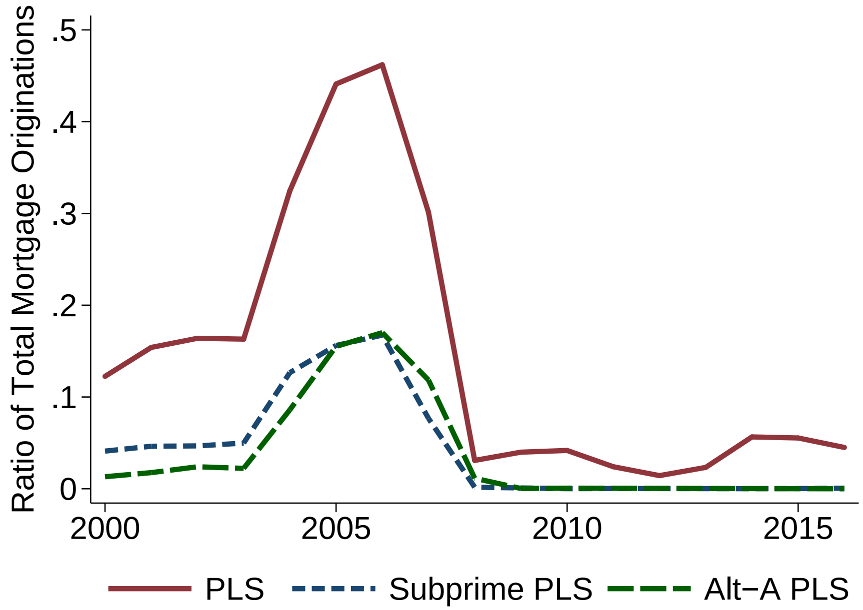
Piskorski, Tomasz, Amit Seru, and James Witkin, "Asset quality misrepresentation by financial intermediaries: evidence from the RMBS market," *The Journal of Finance*, 2015, *70* (6), 2635–2678.

Purnanandam, Amiyatosh, "Originate-to-Distribute Model and the Subprime Mortgage Crisis," *The Review of Financial Studies*, 2011, *24* (6), 1881–1915.

Saiz, Albert, "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

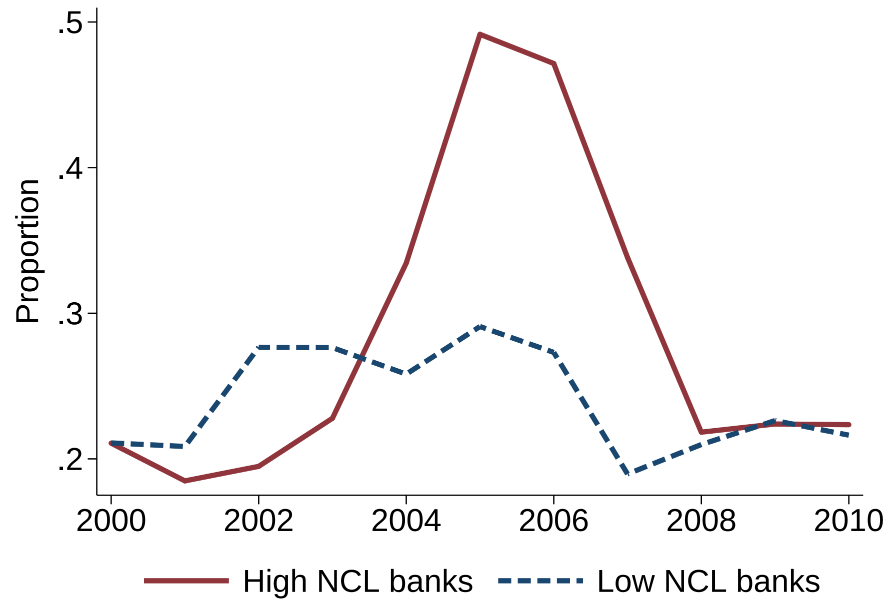
Simsek, Alp, "Belief disagreements and collateral constraints," *Econometrica*, 2013, *81* (1), 1–53.

Figure 1: Acceleration of Private Label Securitization of Mortgages



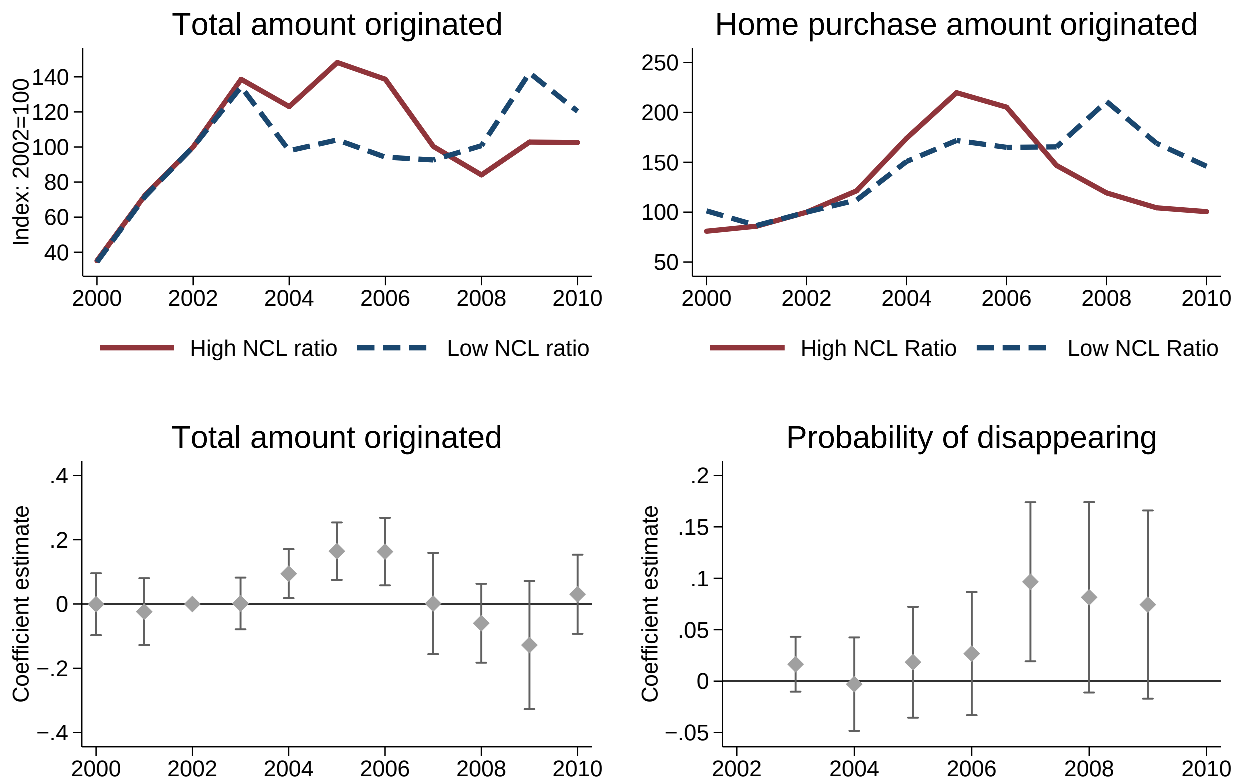
This figure plots the share of total mortgage originations that were sold into private label securitization (PLS), subprime PLS, and Alt-A PLS.

Figure 2: Fraction of Mortgages Sold to Private Institutions by NCL of Originating Bank



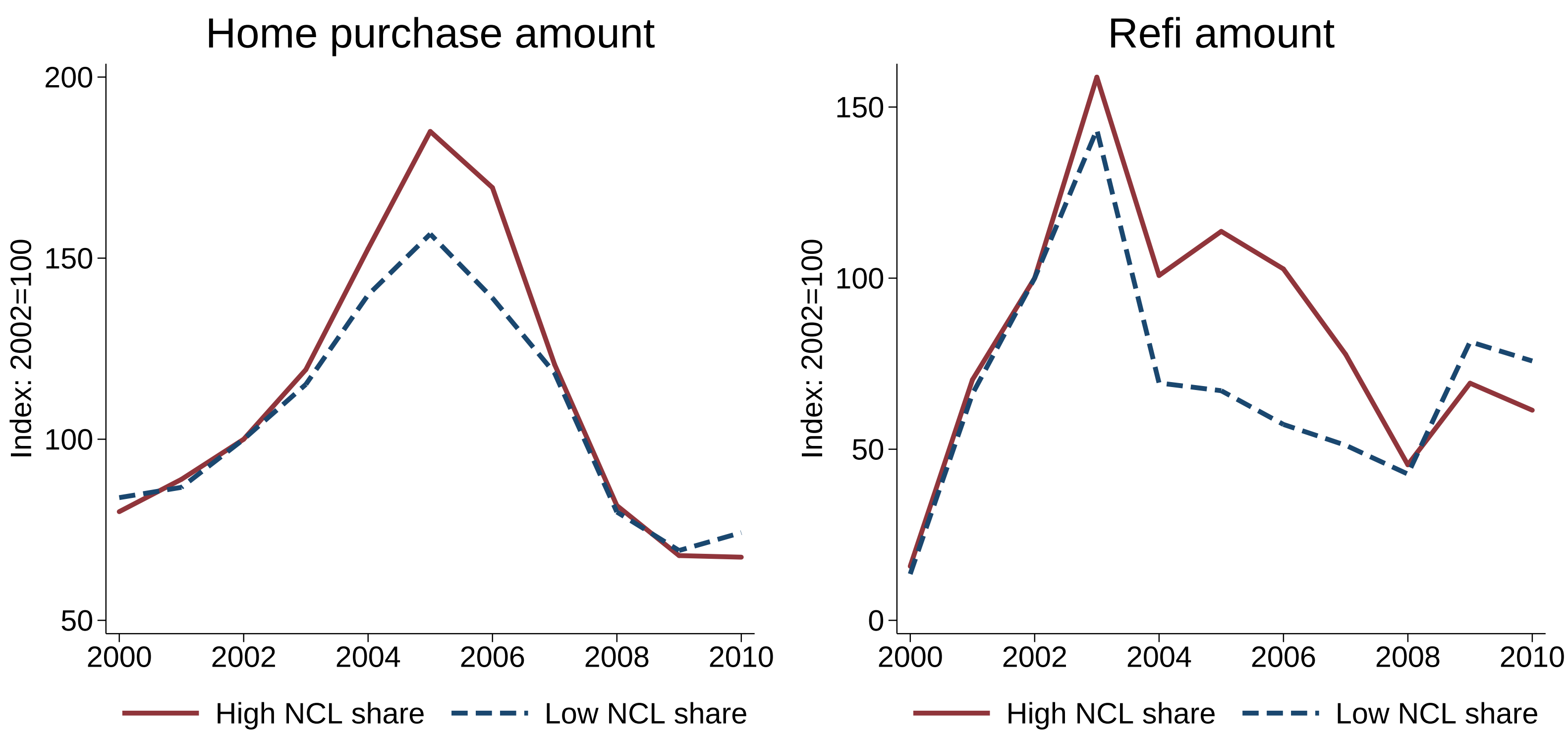
This figure plots the share of total mortgage amounts originated that were sold to private institutions by banks above and below the median non-core liabilities ratio (NCL) as of 2002. NCL is defined as one minus the proportion of liabilities that are federally insured deposits for institutions that are in the FFIEC Call Reports. Non-bank mortgage lenders are excluded.

Figure 3: Relative Expansion of Mortgage Amounts Originated by High NCL Ratio Lenders



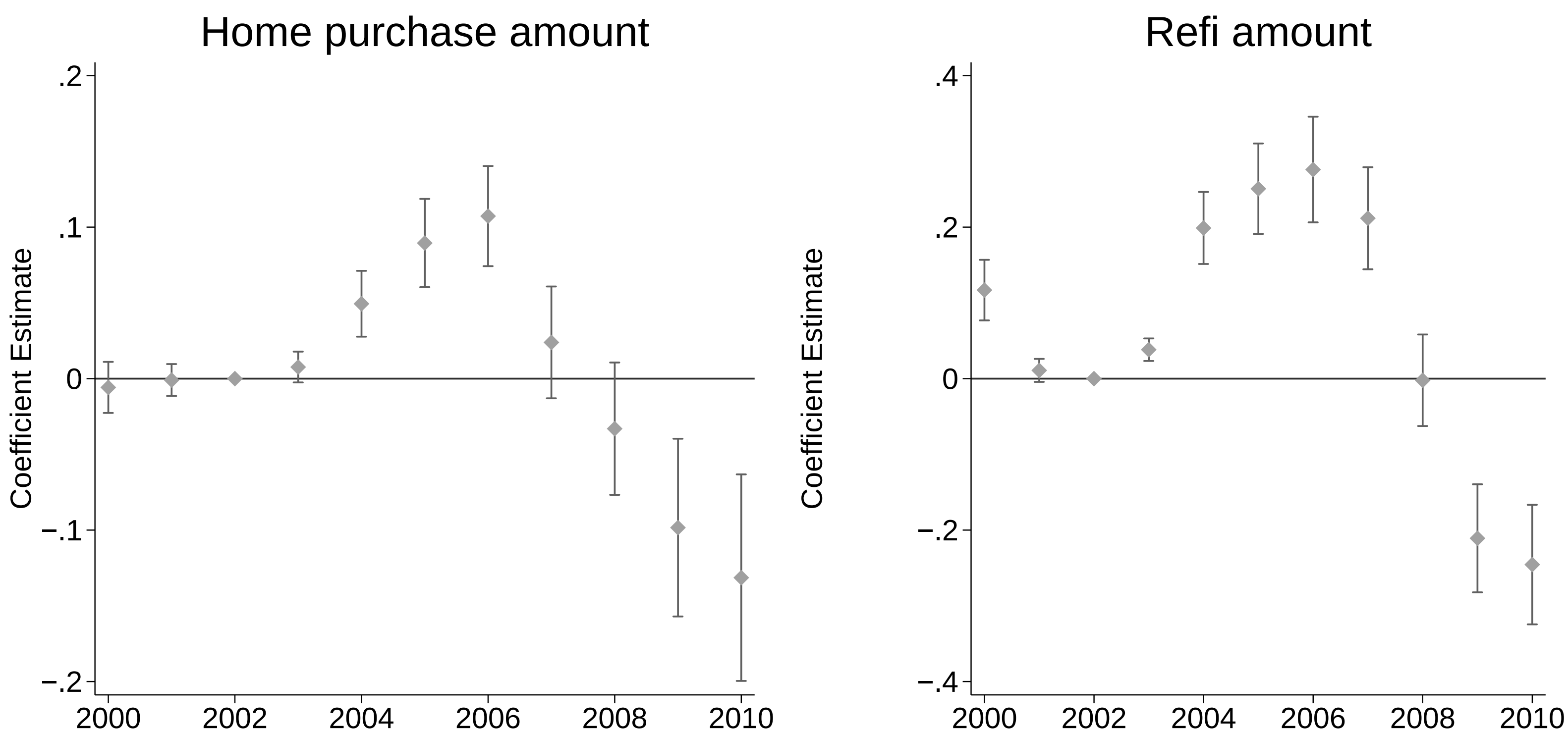
The top left panel plots total mortgage amount originated by lenders above and below the median non-core liabilities ratio (NCL) as of 2002. The top right panel plots home purchase mortgage amount originated by lenders above and below the median NCL as of 2002. The bottom left panel plots the coefficients $\{\beta_k\}$ of the specification $ln(y_{b,t}) = \alpha_b + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{b,2002} + \varepsilon_{b,t}$ for lender b at time t. $y_{b,t}$ is total mortgage amount originated by a lender b in year t. The bottom right panel plots the coefficient $\{\rho_t\}$ of the repeated cross sectional regression $GONE_{b,t} = \alpha + \rho_t NCL_{b,2002} + \varepsilon_b$ where $GONE_{b,t}$ is equal to 1 if a lender in the sample in 2002 is no longer in the sample in year t for years 2003-2009. NCL is defined as one minus the proportion of liabilities that are federally insured deposits for institutions that are in the FFIEC Call Reports and one for institutions regulated by the Department of Housing and Urban Development (HUD). The regressions are weighted by the mortgage amount originated in 2002 by lender b. 95% confidence intervals from robust standard errors are also plotted. Lender fixed effects included in the panel regression.

Figure 4: Mortgage Amount Originated in Zip Codes, by NCL Share of Lenders



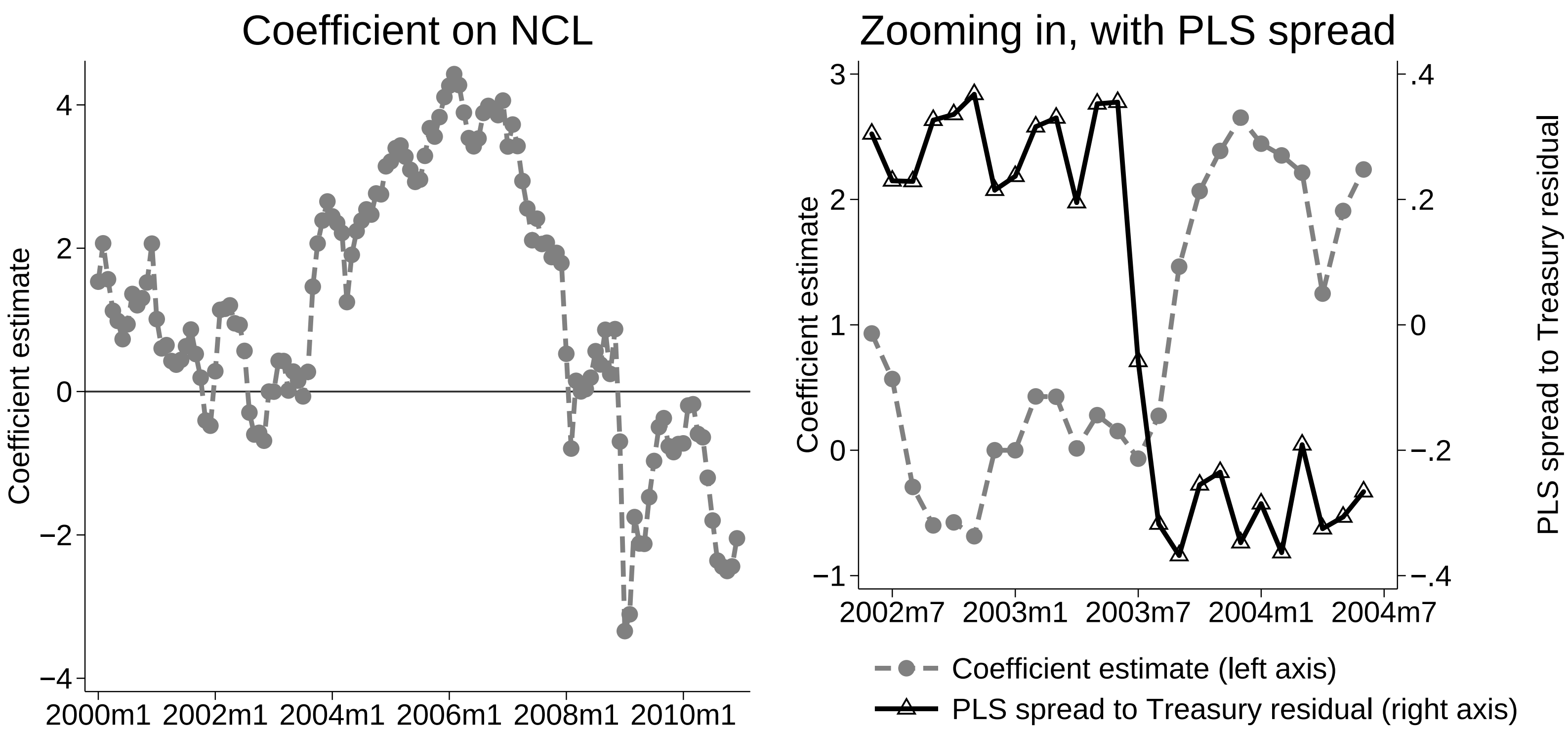
The figure plots the mortgage amount originated in zip codes above and below the median non-core liabilities share (NCL) as of 2002. NCL at the zip code-level is defined as the weighted average of the NCL of lenders where the weights are the share of mortgages originated in 2002 by a lender b in zip code z.

Figure 5: Zip-code Level Mortgage Amounts Originated by NCL Share: Panel Regressions



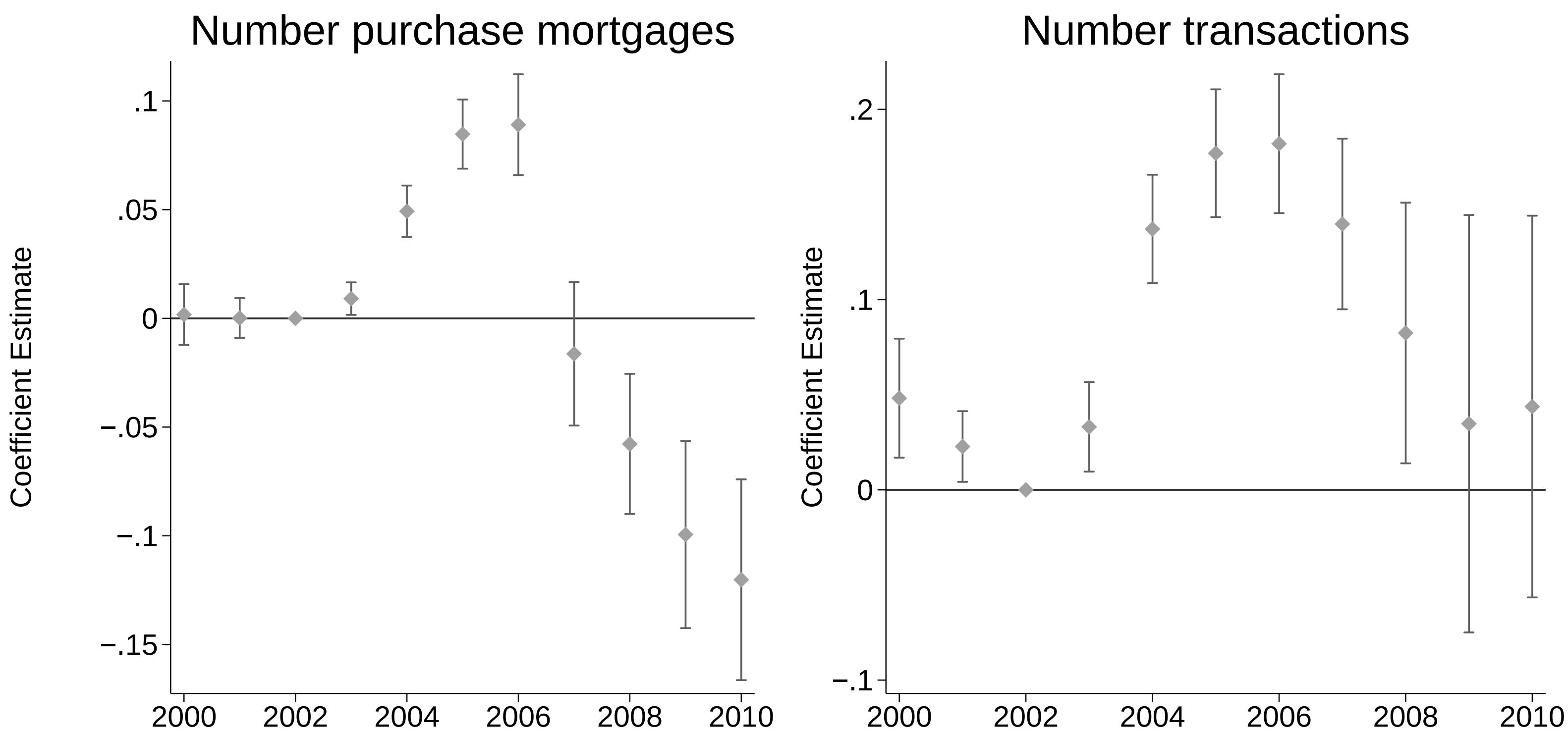
The panels plot the coefficients $\{\beta_k\}$ of the specification $ln(y_{z,t}) = \alpha_z + \gamma_t + \sum_{k\neq 2002} \mathbb{1}_{t=k}\beta_k NCL_{z,2002} + \varepsilon_{z,t}$ for zip code z in year t. $y_{z,t}$ is total mortgage amount originated in zip code z in year t. NCL at the zip code-level is defined as the weighted average of NCL at the lender-level where the weights are the share of loans originated in 2002 by a lender b in zip code z. The regressions are weighted by the share of total occupied housing units in zip code z in 2000. 95% confidence intervals from robust standard errors are also plotted. Zip code level fixed effects included.

Figure 6: MSA Monthly Level Mortgage Amounts Originated by NCL Share: Panel Regressions



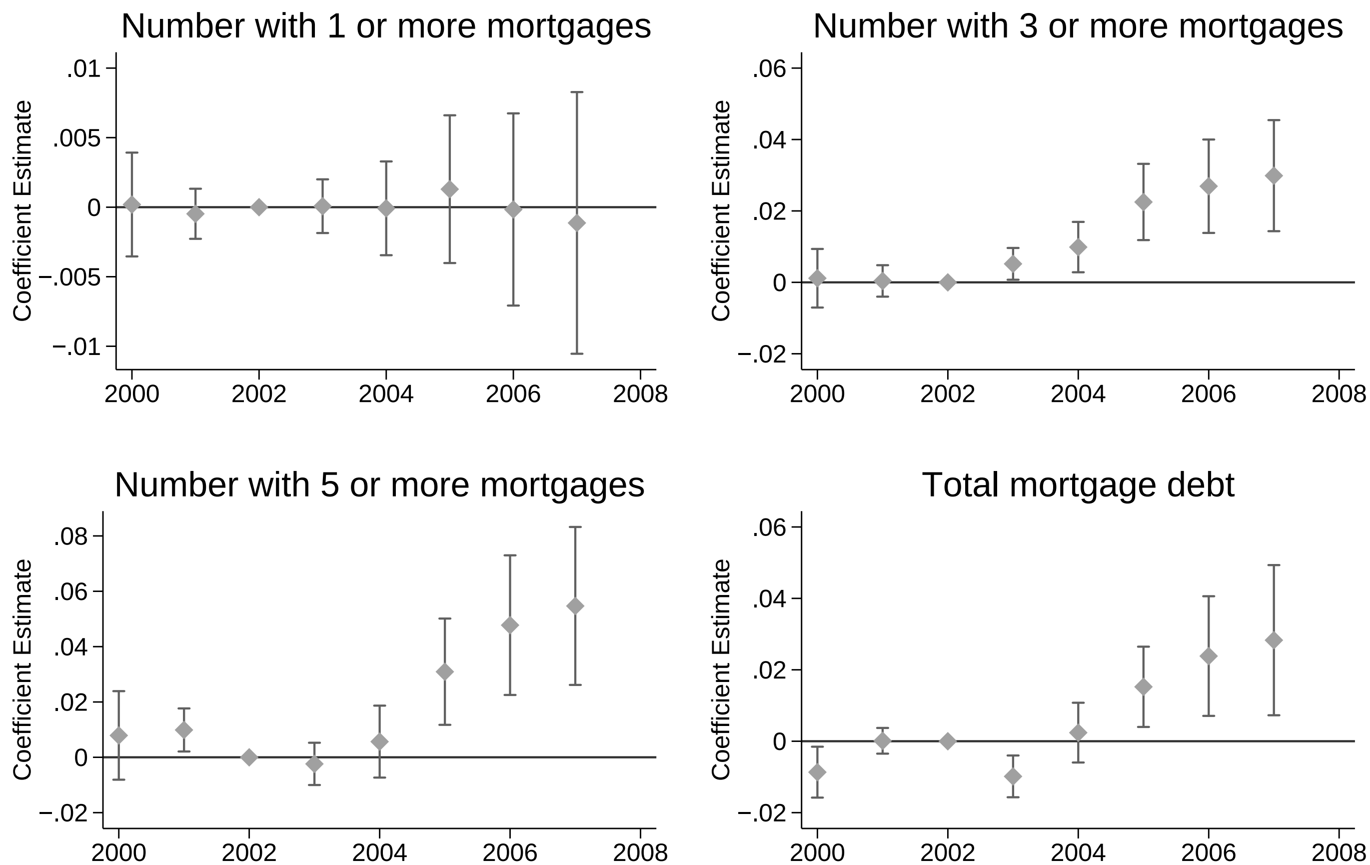
The left panel plots the coefficients $\{\beta_k\}$ of the specification $ln(y_{m,t}) = \alpha_m + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{m,2002} + \varepsilon_{m,t}$ for MSA m at time t. $y_{m,t}$ is total mortgage amount originated in MSA m in year t. The right panel zooms in around 2003 and also plots the PLS mortgage interest spread to Treasury residual from Justiniano et al. (2017). NCL at the MSA-level is defined as the weighted average of NCL at the lender-level where the weights are the share of loans originated in 2002 by a lender b in MSA m. The regressions are weighted by the share of total occupied housing units in MSA m in 2000. 95% confidence intervals from robust standard errors are also plotted. MSA fixed effects included in panel regression.

Figure 7: Zip-code Transaction Volume by NCL Share: Panel Regressions



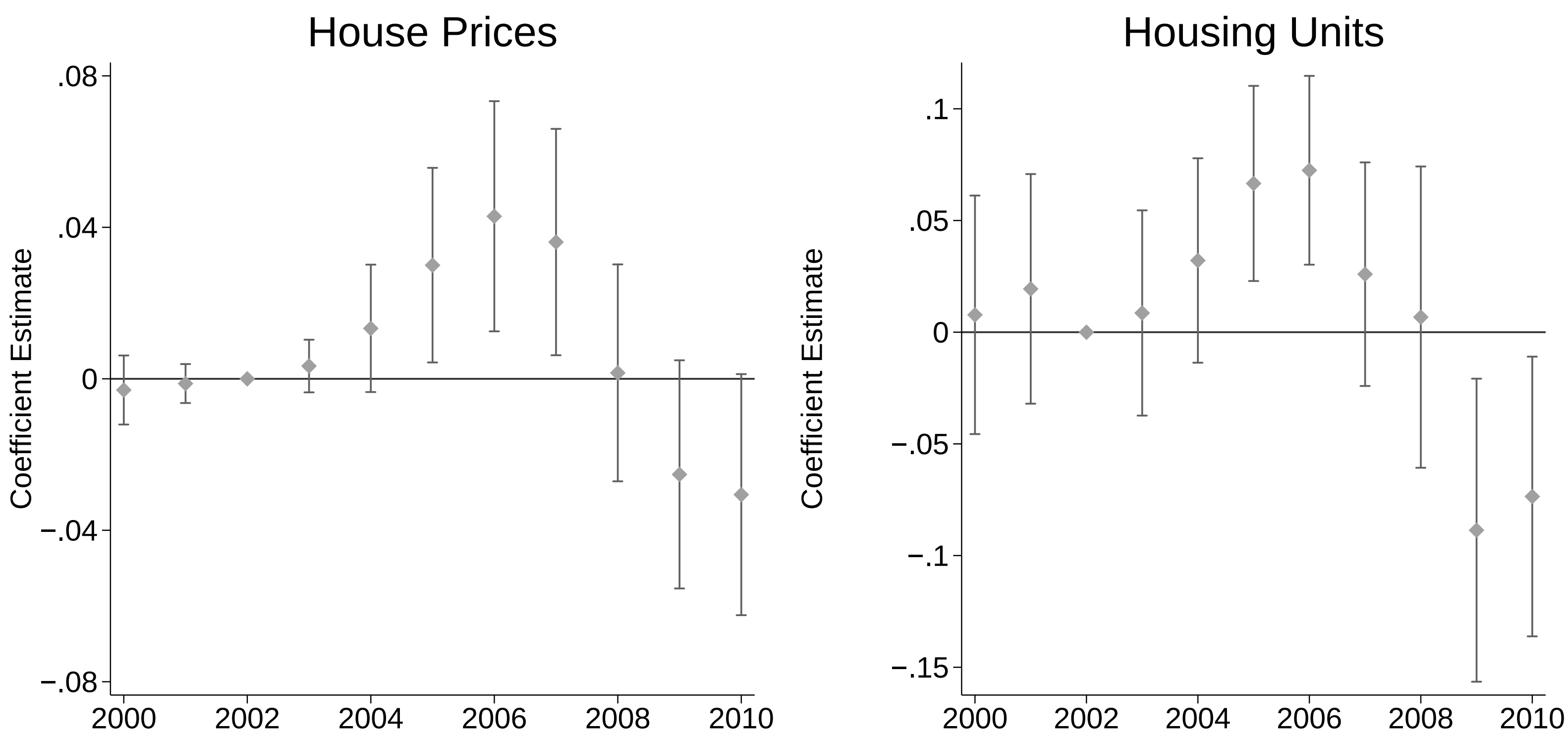
The panels plot the coefficients $\{\beta_k\}$ of the specification $ln(y_{z,t}) = \alpha_z + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{z,2002} + \varepsilon_{z,t}$ for zip code z in year t. $y_{z,t}$ in the left panel is the number of mortgage origination for home purchases. $y_{z,t}$ in the right panel is the number of housing transactions. NCL at the zip code-level is defined as the weighted average of NCL at the lender-level where the weights are the share of loans originated in 2002 by a lender b in zip code z. The regressions are weighted by the share of total occupied housing units in zip code z in 2000. 95% confidence intervals from robust standard errors are also plotted. Zip code level fixed effects included.

Figure 8: Zip-code Growth in Investors by NCL Share: Panel Regressions



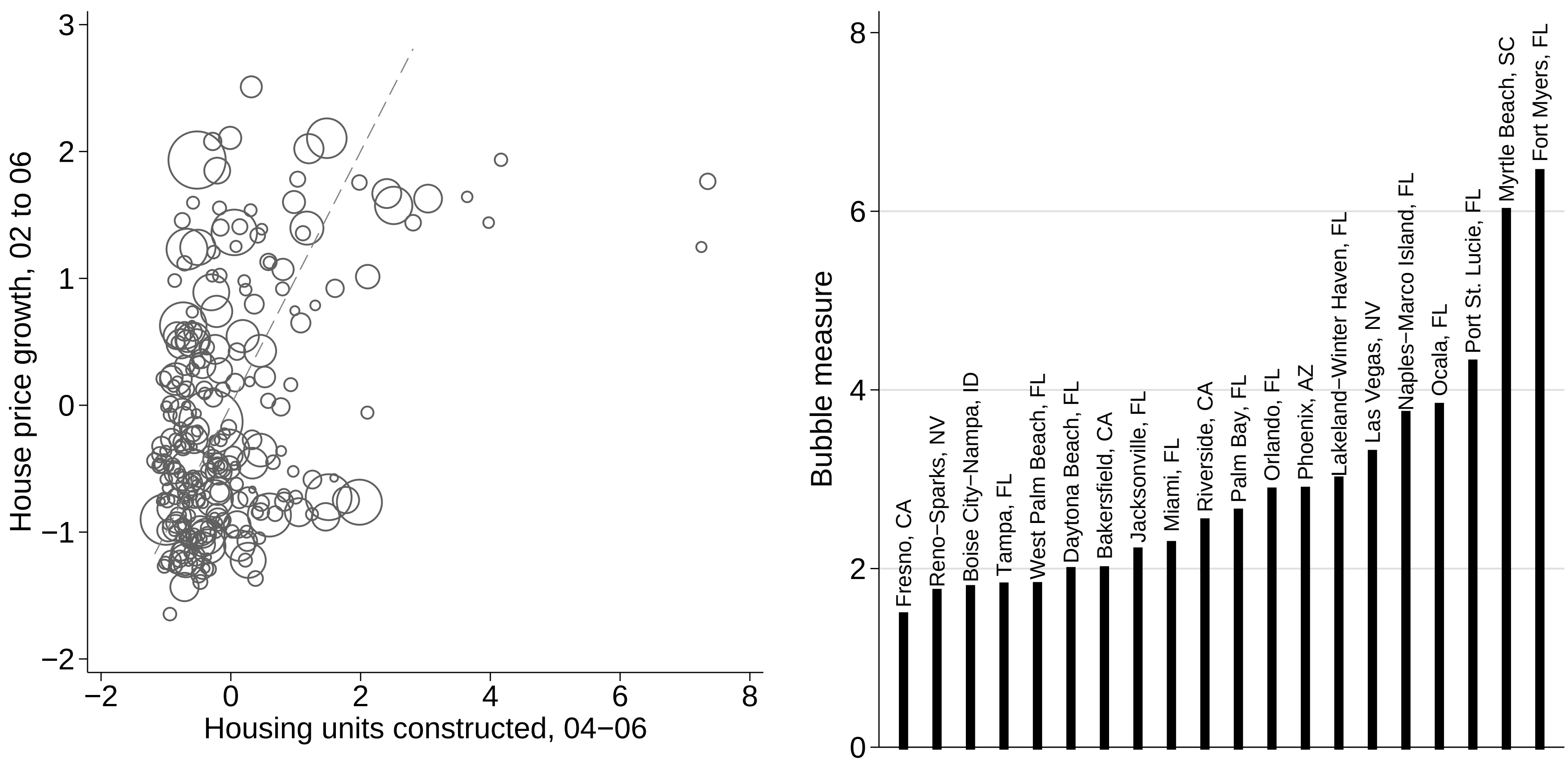
The panels plot the coefficients $\{\beta_k\}$ of the specification $ln(y_{z,t}) = \alpha_z + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{z,2002} + \varepsilon_{z,t}$ for zip code z in year t. $y_{z,t}$ is the number of individuals with 1 or more mortgages in the zip code, 3 or more mortgages, 5 or more mortgages, and total mortgage debt outstanding in the zip code. NCL at the zip code-level is defined as the weighted average of NCL at the lender-level where the weights are the share of loans originated in 2002 by a lender b in zip code z. The regressions are weighted by the share of total occupied housing units in zip code z in 2000. 95% confidence intervals from robust standard errors are also plotted. Zip code level fixed effects included.

Figure 9: House Prices and Construction by NCL share: Panel Regression



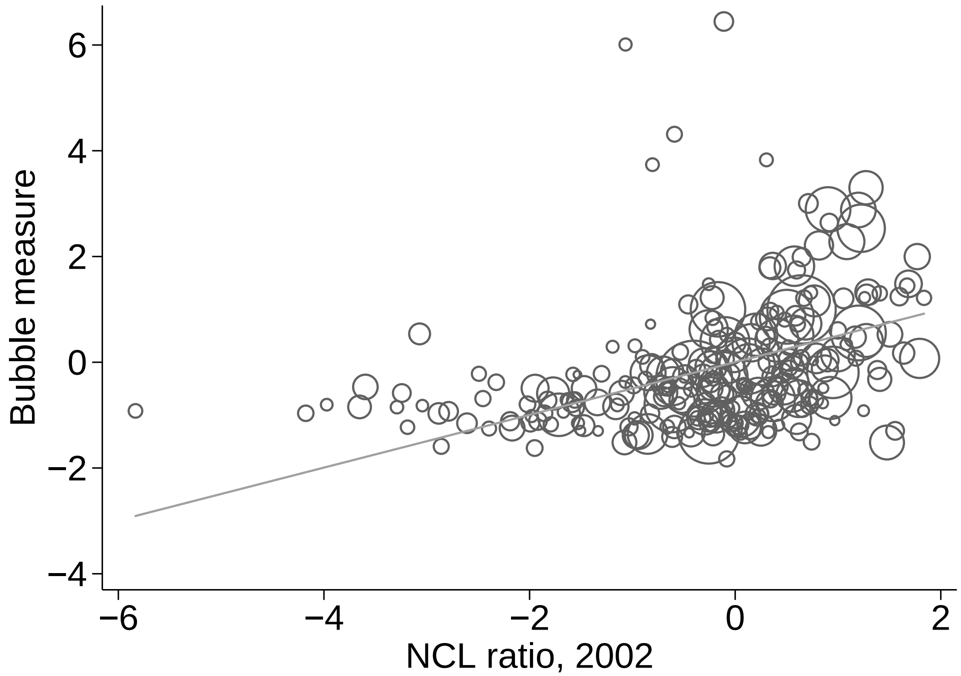
The left panel plots the coefficients $\{\beta_k\}$ of the specification $ln(HP_{z,m,t}) = \alpha_{z,m} + \gamma_t + \sum_{k\neq 2002} \mathbb{1}_{t=k}\beta_k NCL_{z,m,2002} + \sum_{k\neq 2002} \mathbb{1}_{t=k}\beta_k SAIZ_{m,2002} + \varepsilon_{z,m,t}$ for zip code z, MSA m, and year t. The right panel plots the coefficients $\{\beta_k\}$ of the specification $ln(UNITS_{m,t}) = \alpha_m + \gamma_t + \sum_{k\neq 2002} \mathbb{1}_{t=k}\beta_k NCL_{m,2002} + \varepsilon_{m,t}$ for MSA m in year t. $HP_{z,m,t}$ is house prices of zip code z and year t, $SAIZ_m$ is the Saiz (2010) elasticity measure of MSA m, and $UNITS_{m,t}$ is housing units of MSA m and year t. NCL at the zip code-level (MSA-level) is defined as the weighted average of NCL at the lender-level where the weights are the share of loans originated in 2002 by a lender b in zip code z (MSA m). The regressions are weighted by the share of total occupied housing units in their respective geographies in 2000. 95% confidence intervals from robust standard errors are also plotted.

Figure 10: Bubbly Cities: House Price Growth and Construction Growth



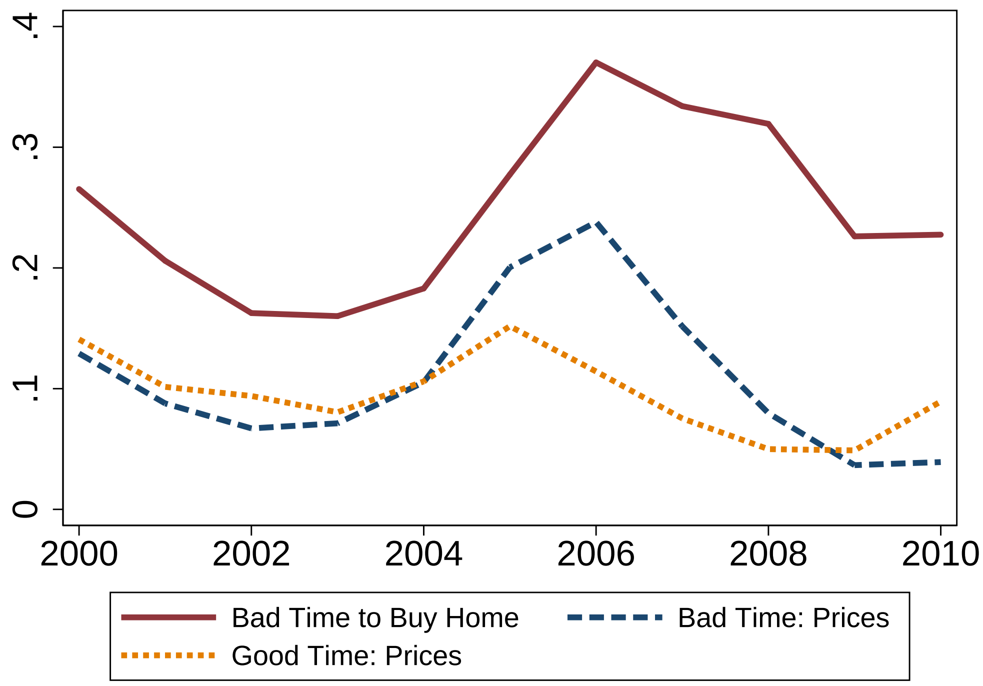
The left panel plots MSA-level house price growth from 2002 to 2006 against construction during the housing boom. The measure of construction is total units constructed from 2004 to 2006 scaled by the total number of units in the MSA as of 2000. Both measures are standardized to be mean zero and standard deviation one. The dashed line is the 45 degree line. The right panel shows the bubble measure for the top 20 bubbly cities. The bubble measure is constructed by drawing a perpendicular line to the 45 degree line in the left panel, and then measuring the distance to the (0,0) point. See text for more details.

Figure 11: NCL Share and Bubbly Cities



This figure plots for each MSA in the sample the bubble measure against the NCL ratio as of 2002. The bubble measure is constructed by drawing a perpendicular line to the 45 degree line in the scatter plot of house price growth against construction growth, and then measuring the distance to the (0,0) point. See text for more details. The NCL ratio is standardized to be mean zero and standard deviation one.

Figure 12: Measures of Optimism on Housing Market from the Michigan Survey



Plots the share of individuals that respond to the question “Good or bad time to buy home” with the answers “good time to buy a home,” “bad time to buy a home,” “bad time to buy because of price considerations,” and “good time to buy because of price considerations”. See the text for more details.

Table 1: Summary Statistics

Panel A: Lender-Level Data

	count	mean	p50	sd
2002 NCL ratio	5026	0.74	0.68	0.20
2002 Non-bank	5040	0.25	0.00	0.43
$\Delta_{02,05}$ PLS share	3950	0.15	0.08	0.26
$\Delta_{00,02}$ PLS share	4130	0.00	-0.02	0.21
$\Delta_{98,00}$ PLS share	3440	0.01	-0.02	0.21
$\Delta_{02,05}$ ln (Amount originated)	3950	-0.02	-0.09	0.73
$\Delta_{00,02}$ ln (Amount originated)	4130	1.22	1.20	0.64
$\Delta_{98,00}$ ln (Amount originated)	3440	-0.02	-0.22	1.16

Panel B: Zip-Code-Level Data

	count	mean	p50	sd
2002 NCL Share	12427	0.77	0.77	0.05
Δ_{BOOM} ln (Amount originated)	12422	0.44	0.38	0.43
Δ_{BOOM} ln (Home purchase amount originated)	12419	0.57	0.54	0.36
Δ_{BOOM} ln (Refinancing amount originated)	12400	0.32	0.23	0.53
Δ_{BOOM} ln (PLS loans originated)	12322	1.65	1.63	0.52
Δ_{BOOM} ln (Number of loans originated for home purchase)	12419	0.29	0.26	0.29
Δ_{BOOM} ln (Volume of housing transactions)	4497	0.32	0.29	0.31
$\Delta_{02,06}$ ln (House Prices)	6619	0.37	0.36	0.22
$\Delta_{02,06}$ ln (Number with 1 or more mortgages)	12418	0.07	0.05	0.12
$\Delta_{02,06}$ ln (Number with 3 or more mortgages)	12403	0.25	0.22	0.17
$\Delta_{02,06}$ ln (Number with 5 or more mortgages)	12254	0.46	0.43	0.27
$\Delta_{02,06}$ ln (Mortgage debt outstanding)	12417	0.46	0.45	0.21

Panel C: MSA-Level Data

	count	mean	p50	sd
2002 NCL Share	259	0.761	0.766	0.039
Housing Supply Elasticity	259	1.958	1.645	1.177
Constructed units, 04-06	259	0.033	0.024	0.027

Table 2: High NCL Ratio Predicts Growth in Mortgage Originations

	Δ Fraction PLS, 02 to 05	Amount growth, 02 to 05			Amount growth, Pre-Boom	
	(1)	(2)	(3)	(4)	(5) 1998-2000	(6) 2000-2002
2002 NCL ratio	0.151**	0.183***		0.203*	-0.006	-0.083
	(0.049)	(0.047)		(0.082)	(0.046)	(0.105)
Non-bank 2002			0.284*	-0.067		
			(0.114)	(0.189)		
Sample	Banks	Full	Full	Full	Full	Full
N	3287.000	3947.000	3950.000	3947.000	3447.000	3433.000
R-sq	0.210	0.061	0.027	0.062	0.000	0.007

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression results for the specification $\Delta PLS_{b,2002,2005} = \alpha + \beta NCL_{b,2002} + \varepsilon_b$ for lender b are in column 1. Regression results for the specification $\Delta HMDA_{b,2002,2005} = \alpha + \beta_0 NCL_{b,2002} + \beta_1 NB_{b,2002} + \varepsilon_{b,2002,2005}$ for lender b are in columns 2 through 4. $\Delta PLS_{b,2002,2005}$ is the change in the share of mortgage amount originated that were sold to a private institution by lender b from 2002 to 2005, $\Delta HMDA_{b,2002,2005}$ is the log change in mortgage amount originated by lender b from 2002 to 2005, and $\Delta HMDA_{b,PRE}$ is the log change in mortgage amount originated by lender b from 1998 to 2000 and 2000 to 2002. NCL is defined here as one minus the proportion of liabilities that are federally insured deposits for institutions that are in the FFIEC Call Reports and one for institutions regulated by the Department of Housing and Urban Development (HUD). A non-bank mortgage lender is an institution regulated by the HUD. Column 1 is restricted to commercial banks and thrifts. Standard errors are robust.

Table 3: NCL Share Correlations with Observatble Variables

Covariates	Across MSA	Within MSA
2000 Deposits/Purchase amount originated	-1.11***	
	(.246)	
Saiz elasticity	-.262***	
	(.071)	
2000 Fraction Vantage below 660	.038***	.075***
	(.007)	(.001)
2000 Fraction age 65+	-.006**	-.010***
	(.001)	(.000)
2000 Fraction hispanic or black	.069***	.110***
	(.011)	(.002)
2000 Fraction renters	.008	.030***
	(.005)	(.001)
2000 Log median home value	.033	-.110***
	(.020)	(.005)
2000 Log median household income	-.009	-.081***
	(.017)	(.003)

Univariate regression coefficients of the non-core liabilities share (NCL) in 2002 to observable variables at the MSA-level (left column) and at the zip code level (right column). The zip-code level regressions include MSA fixed effects, and so we these are within-MSA coefficients. The NCL ratio at the geographical-level is defined as the weighted average of NCL at the lender-level where the weights are the share of loans originated in 2002 by lender b in geography g.

Table 4: High NCL Ratio Predicts Growth in Mortgage Originations: With Geography Fixed Effects

	Bank-MSA Originations		Bank-Zip-Code Originations	
	(1)	(2)	(3)	(4)
2002 NCL Ratio	0.169***	0.140***	0.154***	0.125***
	(0.007)	(0.006)	(0.006)	(0.006)
Geography FE	N	Y	N	Y
N	65446	65446	888272	888272
R-sq	0.041	0.162	0.031	0.204

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression results for the specification $\Delta y_{b,m,2002,2005} = \alpha_m + \beta NCL_{b,2002} + \varepsilon_{b,m,2002,2005}$ for lender b in MSA m are in columns 1 and 2. Regression results for the specification $\Delta y_{b,z,2002,2005} = \alpha_z + \beta NCL_{b,2002} + \varepsilon_{b,z,2002,2005}$ for lender b in zip code z are in columns 3 and 4. $\Delta y_{b,m,2002,2005}$ ($\Delta y_{b,z,2002,2005}$) is defined here as the log change in total mortgage amount originated from 2002 to 2005 for lender b in MSA m (zip code z). NCL is defined here as one minus the proportion of liabilities that are federally insured deposits for lenders that are in the FFIEC Call Reports and one for lenders regulated by the Department of Housing and Urban Development (HUD). Regressions are weighted by the share of loans originated in 2002 by lender b in MSA m (zip code z). Standard errors are clustered at the MSA (zip code) level.

Table 5: NCL Share and Mortgage Origination Growth: Zip Code Level

	Purchase amount		Refi amount		PLS originations		Non-PLS originations	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2002 NCL Share	0.086***	0.117***	0.224***	0.287***	0.128***	0.158***	0.050***	0.100***
	(0.016)	(0.017)	(0.030)	(0.030)	(0.027)	(0.042)	(0.011)	(0.019)
MSA FE	N	Y	N	Y	N	Y	N	Y
N	12419	12419	12400	12400	12322	12322	12235	12235
R-sq	0.056	0.410	0.180	0.670	0.061	0.491	0.021	0.432

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression results for the specification $\Delta y_{z,m,BOOM} = \alpha_m + \beta NCL_{z,m,2002} + \varepsilon_{z,m}$ from zip code z and MSA m. $\Delta y_{z,m,BOOM}$ is defined as the log change in outcome y from 2000-2002 to 2004-2006 in zip code z in MSA m. The outcomes are home purchase mortgage amount originated, refinancing mortgage amount originated, mortgage amount originated into the PLS market, and mortgage amount originated not into the PLS market. The NCL at the zip code-level is defined as the weighted average of NCL at the lender-level where the weights are the share of loans originated in 2002 by a lender b in zip code z. Regressions are weighted by the share of households in zip code z and year 2000. Standard errors are clustered at the MSA-Level.

Table 6: Change in Volume per 2000 Unit: Zip Code Level

	(1) Δ Volume per 2000 unit	(2) Mail address is house address	(3) Mail same zip, different address	(4) Mail outside zip	(5) Mail address missing
2002 NCL Share	0.063***	0.033**	0.002	0.008*	0.019***
	(0.016)	(0.012)	(0.001)	(0.004)	(0.004)
MSA FE	Y	Y	Y	Y	Y
N	4473	4473	4473	4473	4473
R-sq	0.250	0.287	0.250	0.194	0.404

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression results for the specification $\Delta volume_{z,m,BOOM} = \alpha_m + \beta NCL_{z,m,2002} + \varepsilon_{z,m}$ from zip code z and MSA m. $\Delta volume_{z,m,BOOM}$ is the number of housing transactions from 2004 to 2006 minus the number from 2000 to 2002, scaled by the number of housing units in the zip code as of 2000. Columns 2 through 5 split the left hand side variable into transactions where the mailing address of the buyer is (a) the same as the house, (b) in the same zip code but different than the house, (c) is outside the zip code of the house, and (d) is missing. The NCL at the zip code-level is defined as the weighted average of NCL at the lender-level where the weights are the share of loans originated in 2002 by a lender b in zip code z. Regressions are weighted by the share of occupied houselds in zip code z and year 2000. Standard errors are clustered at the MSA-Level.

Table 7: NCL Share and Growth in Number of Investors from 2002 to 2006: Zip Code Level

	1 or more mortgages		3 or more mortgages		5 or more mortgages		Total mortgage debt	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2002 NCL Share	0.015	0.000	0.033*	0.000	0.021**	0.000	0.040***	0.000
	(0.009)	(.)	(0.017)	(.)	(0.008)	(.)	(0.008)	(.)
MSA FE	N	Y	N	Y	N	Y	N	Y
N	385	385	385	385	384	384	384	384
R-sq	0.057	1.000	0.100	1.000	0.030	1.000	0.096	1.000

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression results for the specification $\Delta y_{z,m,2002,2006} = \alpha_m + \beta NCL_{z,m,2002} + \varepsilon_{z,m,2002,2006}$ from zip code z and MSA m. $\Delta y_{z,m,2002,2006}$ is the log change in outcome y in zip code z in msa m from 2002 to 2006. The outcomes are number of individuals with 1 or more mortgages in the zip code, the number with 3 or more mortgages, the number with 5 or more mortgages, and total mortgage debt outstanding in the zip code. The NCL at the zip code-level is defined as the weighted average of NCL at the lender-level where the weights are the share of loans originated in 2002 by a lender b in zip code z. Regressions are weighted by the share of households in zip code z and year 2000. Standard errors are clustered at the MSA-Level.

Table 8: NCL Share and House Price Growth

	House Price Growth, 2002 to 2006					
	(1)	(2)	(3)	(4)	(5)	(6)
Housing Supply Elasticity	-0.127***		-0.122***	0.791***		
	(0.018)		(0.017)	(0.151)		
2002 NCL Share		0.059***	0.043**	0.151***	0.018*	0.063***
		(0.018)	(0.015)	(0.027)	(0.008)	(0.017)
2002 NCL Share X Housing Supply Elasticity				-0.055***		-0.027***
				(0.010)		(0.007)
MSA FE	N	N	N	N	Y	Y
N	5540	5540	5540	5540	5540	5540
R-sq	0.313	0.060	0.345	0.413	0.929	0.933

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression results for the specification $\Delta HP_{z,m,2002,2006} = \alpha + \beta SAIZ_m + \varepsilon_{z,m}$ from zip code z and MSA m is in column 1. Regression results for the specification $\Delta HP_{z,m,2002,2006} = \alpha_m + \beta NCL_{z,m,2002} + \varepsilon_{z,m}$ from zip code z and MSA m are in columns 2 and 4. Regression results for the specification $\Delta HP_{z,m,2002,2006} = \alpha_m + \beta NCL_{z,m,2002} + \beta_1 SAIZ_m + \beta_2 NCL_{z,m,2002} SAIZ_m + \varepsilon_{z,m}$ from zip code z and MSA m are in columns 3 and 5. $HP_{z,m,2002,2006}$ is the log change in house prices from 2002 to 2006 in zip code z and $SAIZ_m$ is the Saiz (2010) elasticity measure of MSA m. The regressions are weighted by the share of total occupied households in zip code z in 2000. Standard errors are clustered at the MSA-Level.

Table 9: NCL Share and Construction

	Housing units constructed, 04-06, scaled by 2000 units				
	(1)	(2)	(3)	(4)	(5)
Housing Supply Elasticity	-0.0012		0.0010	0.0340	0.0303
	(0.0014)		(0.0014)	(0.0217)	(0.0178)
2002 NCL Share		0.0101***	0.0104***	0.0147***	0.0105**
		(0.0018)	(0.0018)	(0.0040)	(0.0036)
2002 NCL Share X Housing Supply Elasticity				-0.0017	-0.0017
				(0.0012)	(0.0010)
Census Division FE	N	N	N	N	Y
N	259	259	259	259	259
R-sq	0.003	0.173	0.175	0.184	0.413

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression results for the specification $UNITS_{m,BOOM} = \alpha + \beta SAIZ_m + \varepsilon_m$ from MSA m is in column 1. Regression results for the specification $UNITS_{m,BOOM} = \alpha + \beta NCL_{m,2002} + \varepsilon_m$ from MSA m is in column 2. Regression results for the specification $UNITS_{m,BOOM} = \alpha + \beta_0 SAIZ_m + \beta_1 NCL_{m,2002} + \varepsilon_m$ from MSA m is in column 3. Regression results for the specification $UNITS_{m,BOOM} = \alpha + \beta_0 SAIZ_m + \beta_1 NCL_{m,2002} + \beta_2 NCL_{m,2002} SAIZ_m + \varepsilon_m$ from MSA m is in column 4. $UNITS_{m,BOOM}$ in MSA m is total units constructed from 2004 to 2006 scaled by the total number of units in the MSA as of 2000. $SAIZ_m$ is the Saiz (2010) elasticity measure of MSA m. The regressions are weighted by the share of total occupied households in MSA m in 2000. Standard errors are robust. Census division fixed effects are included in column 5.

Table 10: NCL Share and Bubbly MSAs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Bubble measure	Bubble measure	Boom city	Boom city	HP growth 02 to 10	HP growth 02 to 10	Δ units 09-11 minus 00-02	Δ units 09-11 minus 00-02
2002 NCL share	0.439***	0.216**	0.079***	0.054**	-0.031*	-0.063***	-0.006***	-0.004***
	(0.078)	(0.069)	(0.023)	(0.020)	(0.012)	(0.013)	(0.001)	(0.001)
Housing supply elasticity	-0.283***	-0.198***	-0.034*	-0.026	-0.017	-0.004	-0.000	0.001
	(0.057)	(0.044)	(0.016)	(0.016)	(0.011)	(0.010)	(0.001)	(0.001)
Census Division FE	N	Y	N	Y	N	Y	N	Y
N	253	253	259	259	253	253	259	259
R-sq	0.290	0.513	0.102	0.243	0.042	0.445	0.130	0.416

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression results are for the specification $y_m = \alpha + \beta_0 SAIZ_m + \beta_1 NCL_{m,2002} + \varepsilon_m$ for MSA m. For columns 1 and 2, y_m is the bubble measure, which is constructed by drawing a perpendicular line to the 45 degree line in the scatter plot of house price growth against construction growth, and then measuring the distance to the (0,0) point. See text for more details. For columns 3 and 4, y_m is an indicator equal to 1 if a MSA is a "boom" city. This is when an MSA's house price growth from 2002 to 2006 is in the top quartile of the distribution, and $UNITS_{m,BOOM}$ is in the top quartile of the distribution. $UNITS_{m,BOOM}$ is total units constructed from 2004 to 2006, scaled by the total number of units in the MSA as of 2000. y_m is the log change in house prices 2002 to 2010 in columns 5 and 6. y_m is the change in units constructed from 2009 to 2011 minus 2000 to 2002, scaled by total number of units in the MSA as of 2000. $SAIZ_m$ is the Saiz (2010) elasticity measure of MSA m. The regressions are weighted by the share of total occupied households in MSA m in 2000. Standard errors are robust.

Table 11: NCL Share and Housing Market Optimism: CBSA-Level

	(1)	(2)	(3)	(4)	(5)	(6)
	$\Delta_{00,02}$ Good time to buy	$\Delta_{00,02}$ Good time to buy bc of prices	Δ_{boom} Bad time to buy	Δ_{boom} Bad time to buy	Δ_{boom} Bad time to buy bc of prices	Δ_{boom} Bad time to buy bc of prices
2002 NCL Share	-0.006	-0.001				
	(0.012)	(0.009)				
HP growth, 02 to 06			0.272***	0.336*	0.337***	0.287*
			(0.049)	(0.157)	(0.051)	(0.119)
Type	OLS	OLS	OLS	IV	OLS	IV
N	237	237	253	253	253	253
R-sq	0.001	0.000	0.196	0.185	0.378	0.369

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression results for the specification $y_m = \alpha + \beta NCL_{m,2002} + \varepsilon_m$ from MSA m are in columns 1-6. The outcomes y_m are the 2000 to 2002 change in the share of respondents answering it is a good time to purchase a home (column 1) and a good time to buy a home because of price considerations (column 2). For columns 3 and 4, y_m is the 2004-2006 average share of respondents saying it is a bad time to purchase a home minus the 2000-2002 average share. For columns 5 and 6, y_m is the 2004-2006 average share of respondents saying it is a bad time to purchase a home because of price considerations minus the 2000-2002 average share. The regressions are weighted by the number of survey participants in an MSA m. In columns 4 and 6, house price growth from 2002 to 2006 is instrumented using the NCL share of the MSA as of 2002. Standard errors are robust.