NBER WORKING PAPER SERIES

CREDIT SUPPLY AND HOUSING SPECULATION

Atif Mian
Amir Sufi

Working Paper 24823
http://www.nber.org/papers/w24823

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2018

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 especially grateful to Neil Bhutta, Nicola Cetorelli, Neale Mahoney, Stefan Nagel, Stavros Peristiani, Tomasz Piskorski, Amit Seru, Philipp Schnabl, and Eric Zwick for help with data access. We received helpful comments from John Griffin, Christopher Palmer, Andrei Shleifer, Eric Zwick, and seminar participants at Chicago Booth, the NBER Corporate Finance meeting, and Princeton University. Tu Cao, Michael Varley, and Hanbin Yang provided excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Credit Supply and Housing Speculation
Atif Mian and Amir Sufi
NBER Working Paper No. 24823
July 2018
JEL No. E03,E44,G01,G12,G2,G4,R21,R31

ABSTRACT

Credit supply expansion fuels housing speculation, generating a boom and bust in house prices. U.S. zip codes more exposed to the 2003 acceleration of the private label mortgage securitization (PLS) market witnessed a sudden and large increase in mortgage originations and house prices from 2003 to 2006, followed by a collapse in house prices from 2006 to 2010. During the boom, cities with higher PLS-market exposure were more likely to see a large increase in house prices despite substantial new construction; these cities experienced a severe bust after 2006. Most of the marginal home-buyers brought into the housing market by the acceleration of the PLS market were short-term buyers or ``flippers." These marginal buyers had lower credit scores and higher ex post default rates. Speculation by such home-buyers contributed to a large rise in transaction volume from 2003 to 2006, and helped trigger the mortgage default crisis in 2007.

Atif Mian
Princeton University
Bendheim Center For Finance
26 Prospect Avenue
Princeton, NJ 08540
and NBER
atif@princeton.edu

Amir Sufi
University of Chicago
Booth School of Business
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
amir.sufi@chicagobooth.edu

Charles P. Kindleberger wrote that "asset price bubbles depend on the growth in credit" (Kindleberger and Aliber (2005)). There is a long tradition of models exploring how credit can affect asset prices. In Allen and Gorton (1993) and Allen and Gale (2000), easy credit encourages speculators to pay more than the fundamental value of an asset because they can shift downside risk to lenders. In models with asymmetric information and heterogeneous beliefs, price dynamics evolve in which speculators attempt to "ride the bubble" in order to sell the asset to a "greater fool" (Miller (1977), Harrison and Kreps (1978), Scheinkman and Xiong (2003), Abreu and Brunnermeier (2003)); greater availability of credit in such models can fuel this process by allowing more irrational traders to enter the market. Credit can also boost the buying power of more optimistic agents, thereby increasing asset prices. If such optimists default and are forced to sell, asset prices can fall considerably as pessimists become the marginal buyers (Geanakoplos (2010), Simsek (2013)).

While there is long history of models exploring how credit-driven speculation can fuel a boom and bust in asset prices, empirical tests of this idea have proven difficult. Such a test necessitates both an exogenous shock to the supply of credit and the ability to track the characteristics of marginal buyers who are brought into the asset market by the shock.

This study attempts to meet both challenges. First, it identifies a natural experiment–the global rise of shadow banking and private label securitization–to generate plausibly exogenous local variation in credit supply expansion. Second, it utilizes a novel data set covering housing transactions, information on housing market optimism, and the characteristics of individuals obtaining new mortgages. The data set allows for the estimation of the effect of credit expansion on house prices, volume, and speculative trading activity.

The setting is the acceleration of the private-label mortgage backed securitization market (the PLS market) in the fall of 2003.[1] As the left panel of Figure 1 shows, there was a significant rise in the fraction of originated mortgages sold into the PLS market after the

[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.

2000 to 2002 refinancing wave. By 2006, almost half of new dollars originated were sold into the PLS market. The right panel of Figure 1 shows the PLS mortgage interest spread relative to Treasuries controlling for risk and other contractual features from Justiniano et al. (2017). The spread collapsed in the late summer of 2003. The simultaneous rise in quantity and decline in interest spreads is the hallmark of a credit supply expansion, which was related to the global rise of shadow banking during this time period (e.g., Gorton and Metrick (2010), Gennaioli et al. (2012), Cetorelli and Peristiani (2012), Pozsar et al. (2013)).

The empirical strategy is based on the idea that the acceleration of the global shadow banking market in 2003 disproportionately reduced the cost of financing for lenders that traditionally relied on non-core deposit liability financing (high NCL lenders hereafter). These lenders boosted mortgage lending suddenly in 2003. Zip codes more exposed to these lenders as of 2002 were more exposed to the acceleration of the PLS market. They witnessed a sudden rise in mortgage originations at the end of 2003 that corresponds exactly to the sudden decline in PLS mortgage interest spreads shown above in Figure 1. There are a number of results that support the identification assumption that the surge in credit availability in these zip codes was due to exposure to high NCL lenders as opposed to alternative factors. For example, high NCL lenders boosted mortgage origination growth by more than low NCL lenders, and the difference is similar if only within-zip code variation is utilized.

Using this plausibly exogenous source of variation in mortgage origination growth across U.S. zip codes, the analysis shows a significant independent effect of the acceleration of the PLS market on the boom and bust in house prices from 2002 to 2010. In particular, zip codes with higher ex ante exposure to the PLS market witnessed a significantly more amplified housing boom and bust that corresponded in timing with the credit expansion and contraction.

The results suggest that the PLS-market-driven credit expansion fueled a “bubble” dynamic in house prices. Cross-sectional exposure to the PLS market known as of 2006 predicts the subsequent decline in house prices from 2006 to 2010. This predictability is both large in magnitude and statistically precise. Further, credit expansion fueled by the PLS market boosted house prices even in cities such as Las Vegas and Phoenix that experienced substantial new construction during the boom. The phenomenon of high house price growth in such cities has puzzled researchers given that in standard models the ability to cheaply construct more housing units should put a lid on house price growth (e.g., Glaeser et al. (2008), Davidoff (2013), Nathanson and Zwick (2017)). The simultaneous rise in house prices and construction activity led to a severe bust: cities with high PLS market exposure saw a decline in house prices and construction activity that was so severe that by 2010 they found themselves below pre-2002 levels.

Credit-fueled speculation appears to be the critical channel through which the acceleration of the PLS market led to a boom and bust in house prices. For example, zip codes more exposed to the PLS market saw a sudden and large relative rise in transaction volume beginning in 2003 just as mortgage origination growth accelerated. All of the relative rise in volume in high PLS-exposure zip codes was driven by a rise in homes purchased with a mortgage, highlighting the importance of credit in explaining the increase in trading activity.

A novel mortgage-level data set from TransUnion allows for measurement of the characteristics of the marginal borrowers driving the rise in volume. "Flippers," or individuals who buy and sell properties quickly, were a large part of the story. In fact, almost 80% of the relative rise in volume in zip codes with more exposure to the PLS market was driven by flippers. In addition to being predominately associated with flipping, the marginal homebuyers brought into the housing market by the acceleration of the PLS market had lower average credit scores, were younger, and defaulted ex post at significantly higher rates.

These findings point to a small group of speculators as being responsible for the large relative increase in volume in high PLS exposure zip codes. For example, during 2005 and 2006, less than 1% of individuals in the TransUnion data set flipped houses. Results using city-level data on the evolution of housing market optimism from the Michigan survey supports this interpretation. On average, the population in high-PLS exposure cities actually became more *pessimistic* about the housing market from 2003 to 2006 as prices rose. The latter finding is difficult to reconcile with models in which a common housing optimism shock to all individuals in the economy led to a rise in house prices. Instead, together with the results on marginal buyers, the findings support models in which increased credit availability transfers buying power to a small group of "optimists" or "speculators" who can have large effects on volume and prices in housing markets (e.g., Piazzesi and Schneider (2009), Geanakoplos (2010), Simsek (2013), Burnside et al. (2016), DeFusco et al. (2018)).

The dynamics of the crash starting in 2007 are also consistent with these models. Flippers brought in by the PLS market experienced extremely high default rates that began to rise as early as 2006. By the end of 2007, well before the heart of the recession, flippers in zip codes most exposed to the PLS market had a default rate of almost 20%, and it climbed to 35% by 2009. As prices crashed and leveraged individuals receded, the fraction of purchases by individuals not using a recorded mortgage in HMDA increased substantially. Marginal buyers during the boom used substantial leverage, whereas marginal buyers during the bust were more likely to use cash. This is consistent with models such as Geanakoplos (2010) in which prices drop significantly because "optimists" with higher asset valution recede and "pessimists" with lower asset valuation become the marginal buyers.

There is a large body of related research on housing speculation, credit supply, the PLS market, and the housing boom and bust from 2000 to 2010. Given the large body of research, we first present our results and then discuss how our findings are related to the existing literature in Section 6.

1 Data and Summary Statistics

1.1 Data

The main data sets used in this study are at the individual, mortgage, lender, MSA, and zip code level. The two main sources of data are the Home Mortgage Disclosure Act (HMDA) and TransUnion. The HMDA data set records the universe of mortgage originations for mortgage originators that have an office within metropolitan statistical areas (MSAs).[2] 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 the Federal Reserve as a deposit-taking institution. Furthermore, we link these financial institutions to Call Report data using a key provided to us by the Federal Reserve Board. The non-core liability ratio (NCL ratio) is defined to be one minus the ratio of core deposits to total liabilities for banks. It is defined to be one for non-banks that do not take deposits. Please see the appendix and Section 2 for more details.

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 that are located within metropolitan statistical areas. 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 share). The 2002 NCL share 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.

There are two additions to the standard HMDA data. First, we use an MSA by month level version of the HMDA data set below in some specifications. Second, for home purchase mortgages, the HMDA data split first- and second-lien mortgages beginning in 2004. For the years prior to 2004, we use data from Bhutta and Keys (2018) that split first- and second-liens based on a methodology explained in their study.[3]

The analysis also uses mortgage and individual-level credit bureau data from TransUnion, available through the Kilts Center at Chicago Booth. The available TransUnion data are

[2]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."

[3]We are extremely grateful to Neil Bhutta who provided us access to the key linking the HMDA Report ID's to the TFR bank ID's and to the two additions to the standard HMDA data.

a 10% random sample of the universe, but we use only a 5% of this 10% sample given computation limitations, for a 0.5% sample of the universe. From this data set, we construct a mortgage origination data set that is similar to the HMDA data, but it has the advantage of having information on the credit score of the individual, the age of the individual, how long the mortgage account was open, the delinquency status of mortgages taken on by the individual, and whether the individual taking out the mortgage has taken out multiple other mortgages around the same time.

The TransUnion data do not contain any explicit flag for whether a mortgage is a refinancing, a home purchase loan, or a first-lien. In the appendix, we describe how we classify mortgages into these categories, and we provide evidence that our methodology produces aggregate statistics in line with HMDA (see Appendix Figure 1). We also describe how we assign a zip code of the house being purchased to an individual in the TransUnion data. As shown below, all coefficients that can be estimated with both the HMDA data and TransUnion data are remarkably similar regardless of the data source.

Measuring housing market optimism prior to 2007 is a challenge, as large sample data sets on beliefs about the evolution of house prices and the housing market prior to 2007 are unavailable.[4] We follow Piazzesi and Schneider (2009) and use the Michigan Survey of Consumers to measure beliefs about the housing market. Using these questions, we construct a MSA-year level variables on the fraction of respondents saying now is a good time or bad time to buy a house, and also the fraction saying that now is a good or bad time to buy a house because of price or credit considerations. Please see the appendix for more details.

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 CoreLogic, a private vendor that collects and standardizes publicly available

[4]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.

tax assessments and deed records from municipalities across the United States. We use the DeFusco et al. (2018) version of this data set, which is filtered to get the most accurate measure of volume at the zip code level.[5]

1.2 Summary statistics

Table 1 provides summary statistics for the lender-level, zip code-level, mortgage-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. Growth in home purchase and refinancing amount originated during the boom was on average 57% and 32%, respectively. The average growth in the number of first-lien purchase mortgages and growth in transaction volume was very similar, suggesting that the volume boom was associated with individuals taking out mortgages.

The average first-lien purchase mortgage in the TransUnion data set is $188 thousand. A significant advantage of the TransUnion data is that it allows us to measure detailed characteristics of individuals taking out mortgages. For example, the average VantageScore of an individual taking out a first-lien mortgage between 2001 and 2010 is 692, which is at the highest end of the near-prime category. The average age at origination is 42 years, and 24% of individuals that take out a first-lien mortgage default on a mortgage at some point in 2006 or after. As shown below, the TransUnion data allows for the classification of borrowers into a variety of categories, such as "flippers." 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.

[5]We are extremely grateful to the authors for sharing the zip-year level version of their total volume variable. Please see the appendix of DeFusco et al. (2018) for more information on the data construction.

2 Empirical Strategy

2.1 The rise of shadow banks and 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 (2010), Gorton and Metrick (2012), Gorton and Metrick (2013), and Pozsar et al. (2013)), and was likely not due to specific views on U.S. housing markets. For example, Appendix Figure 2 shows that originated amounts in the collateralized loan obligation market, which focuses exclusively on corporate debt with no direct link to residential mortgages, increased from less than \$20 billion to almost \$90 billion from 2002 to 2006. The empirical strategy does not take a stand on the precise source of the aggregate credit supply shock driving the rise of securitization and shadow banking during this time-frame. Researchers have put forth a number of explanations including a global savings glut (Bernanke (2005)), neglected risks by investors (Gennaioli et al. (2012)), or lower uncertainty.

The specific timing of when the PLS market accelerated is shown in Justiniano et al. (2017). They show that the Federal Reserve signaled an end to the easing cycle in the summer of 2003, which led to a rise in longer-term Treasury rates and a collapse in mortgage refinancing for conforming GSE-backed mortgages. Justiniano et al. (2017) argue that mortgage originations for the PLS market accelerated directly during this period, and mortgage interest spreads over Treasuries fell sharply (see Figure 1). As shown below, this is the same time period in which lenders that relied on non-core deposits in their liability structure began expanding credit more than other lenders. Xiao (2018) provides a potential explanation for why shadow banking in general accelerated when the Federal Reserve first signaled an end to lower interest rates. In his model, as the Federal Reserve tightens monetary policy, interest rates on shadow banking liabilities relative to traditional deposits rise, which increases the relative attractiveness of shadow banking liabilities (see also Drechsler et al. (2017) and Nagel (2016)). The aggregate shadow banking pattern is consistent with this observation, as asset-backed securitization increased across a large number of asset classes from 2003 to 2006.

2.2 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. Furthermore, 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 cross-sectional approach used here is based on the idea 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.[6]

An assumption underlying the strategy is that it is difficult for lenders to alter their liability structure, at least at the frequency of "shadow banking" cycles. This assumption is supported by the fact that there is a large degree of persistence in the ranking of banks by the NCL share during the boom. From 2001 to 2007, a regression of the ranking as of 2002 on the ranking as of 2001 through 2007 yields a coefficient between 0.9 and 1 for every year.

2.3 Growth in mortgage originations: lender-level specifications

The identification assumption behind the empirical strategy is that the global rise in shadow banking passes through more directly into a boost in lending by high NCL banks. As a first test of this assumption, 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 outcome variable comes from 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) who group together five categories in the HMDA data that are a rough measure of mortgages sold into the PLS market.[7]

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

[6]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.

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

amount sold to a private institution from 2002 to 2005.

In columns 2 through 4, we present the 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 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 2 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. 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.

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.[8]

2.4 Geographic exposure to High NCL lenders

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 the sample, the 2002 NCL share is calculated as the average of the 2002

[8] In Appendix Figure 3, 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.

NCL ratios of mortgage lenders in the area, where the average is weighted by the amount of mortgage originations in 2002.

High NCL lenders focus on particular markets, which is made clear by characteristics of zip codes into which high NCL lenders originated mortgages as of 2002. Table 3 presents univariate regression coefficients for a set of observable variables regressed on the NCL share 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. This is 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 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.

Furthermore, there is a great deal of persistence in the zip codes serviced by high NCL lenders. As Table 3 shows, both the across- and within-MSA correlation between the 1998 NCL share of the zip code and the 2002 NCL share of the zip code is 0.84 with a small standard error. As we show in the appendix, all main results are robust to the use of the 1998 NCL share of the zip code instead of the 2002 NCL share as the main right hand side variable. High NCL lenders focus where they focus, and this does not appear to change much over time.

2.5 Shocks other than credit supply?

The fact that high NCL lenders focus on zip codes with particular characteristics raises the concern that alternative shocks in these zip codes other than the acceleration of the PLS market may be responsible for the patterns shown below. There are five results that mitigate such omitted variable bias concerns: (1) high NCL lenders expand lending by more than low NCL lenders from 2002 to 2006, and this result is similar even if only within-zip code variation is used, (2) there is a lack of pre-trends in the outcome variables of interest, (3) there is no evidence of a differential housing market optimism shock in high NCL areas from 2000 to 2002, (4) the zip-code level results are robust to the inclusion of MSA fixed effects, and (5) the timing of the relative expansion of mortgage credit in the fall of 2003 in high NCL share areas corresponds exactly with the acceleration of the PLS market. All five results are shown in the analysis below.

The first result relies on lender-MSA and lender-zip code level data sets, which allow for the inclusion of 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 2 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. The estimate in column 4 implies that high NCL lenders expanded lending by more than low NCL lenders, and the relative growth is similar even when evaluating lending by high and low NCL lenders within the same zip code. This specification non-parametrically absorbs any zip-code level shocks, and it shows that high NCL lenders boost lending significantly more than low NCL lenders.[9]

2.6 Credit supply expansion in high NCL share zip codes

Having shown the expansion in credit supply at the lender level, this sub-section turns to the zip-code level analysis. Figure 3 presents coefficients from the following specification:

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

The left hand side variables are the natural logarithm of home purchase mortgage origination amount (left panel) and refinancing mortgage origination amount (right panel) 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

[9] In fact, this specification is likely an example of overly strong controls. 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.

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. There is no evidence that the expansion began before 2003.

Figure 4 presents coefficients for a specification similar to equation 2, except it uses data at a monthly frequency to isolate the exact timing of the credit expansion. This data set is available at the MSA by month level, and the outcome variable is total mortgage amounts originated. As shown in the figure, 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), shown above in Figure 1. 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; it is unlikely that income prospects or housing market optimism increased by more in high NCL share MSAs suddenly in August, September, and October of 2003.

Columns 1 and 2 of Table 5 presents regression coefficients from the following specification:

$$\Delta optimism_{m,00-02} = \alpha + \beta NCL_{m,2002} + \varepsilon_m$$

where $optimism_{m,00-02}$ is a measure of the change in optimism on the housing market from 2000 to 2002 in MSA m. These measures are from the Michigan survey. There is no evidence of a change in housing market optimism prior to the sudden expansion of mortgage lending in 2003, and the coefficient estimates have small standard errors.

Columns 3 through 6 of 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 columns 4 and 6, 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 boost the R^2 by at least a factor of four across the specifications, and their inclusion is increasing 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 implies 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 in refinancing amount originated.

3 The Boom and Bust in House Prices

3.1 House prices

The sudden acceleration of the PLS market at the end of 2003 boosted mortgage originations in high NCL share zip codes, which led to a boom and bust in house prices. Figure 5 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 PLS market collapsed in 2007, which corresponds to a collapse in house prices in high NCL share zip codes. In fact, the collapse was severe enough that the log house price level ended up lower in 2009 and 2010 than its 2002 level relative to low NCL share zip codes. We will return to this point below.

Table 6 presents regression estimates from the following equation:

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

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 with MSA fixed effects.

As the estimate in column 1 shows, a higher 2002 NCL share in a zip code predicts higher 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 3 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 columns 4 and 5, we include MSA fixed effects in order to examine only within-MSA zip code level variation in house prices.[10] 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 MSA. Furthermore, 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 1 and 3, the R^2 increases from 0.06 to 0.93. Such a large increase in the R^2 suggests that CoreLogic is smoothing its zip code level house price indices within MSAs. Such smoothing in the 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 for a decline in the coefficient estimate 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.

Credit expansion from 2002 to 2006 also predicts a larger subsequent decline in house prices from 2006 to 2010, as shown in columns 6 and 7. A one standard deviation increase in

[10] 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.

the NCL share as of 2002 predicts an 8% relative decline in house prices from 2006 to 2010. The specification with MSA fixed effects indicates a 5% relative decline. It is important to emphasize that columns 6 and 7 utilize a measure of exposure to the credit expansion as of 2002, which predicts a stronger decline in house prices from 2006 to 2010. The ability of an ex ante credit expansion to predict a subsequent decline in asset prices echoes findings in Baron and Xiong (2017), who find that credit expansion systematically predicts a crash in bank equity prices in a panel of 20 countries from 1920 to 2012.

3.2 Bubble MSAs

In the cross-section of zip codes, a measure of exposure to credit expansion known prior to 2006 predicts a relative decline in house prices from 2006 to 2009. This suggests that credit expansion fueled a "bubble" in house prices. To investigate this result further, we follow the logic of Glaeser et al. (2008) who show that in a model with rational agents, it is difficult to generate a large rise in house prices in areas in which construction of new homes is cheap. Instead, they show that the presence of irrational home-buyers can generate substantial price increases even if construction activity responds aggressively.[11]

To explore this idea, we define a bubble MSA as one that experiences a large rise in house prices from 2002 to 2006 despite strong new construction activity.[12] The left panel of Figure 6 plots house price growth from 2002 to 2006 against construction activity during the boom for the MSAs in the sample. Construction activity is measured as total new units constructed from 2004 to 2006, scaled by total units as of 2000. Both measures are then 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 rotating the axis counter-clockwise by 45 degrees and then measuring the vertical distance on the rotated graph. This measures the degree to which an MSA witnessed both a large rise in house prices and strong construction

[11] See also Davidoff (2013) and Nathanson and Zwick (2017). The latter study argues that such a pattern could be consistent with a model of speculation.

[12] The MSA is the unit of observation in this section because new construction data from the Census is not available at the zip code level.

activity during the boom.

The right panel shows the top 20 MSAs in terms of the bubble measure. Notable examples include Orlando and Phoenix, two cities emphasized in Glaeser et al. (2008). The use of the term bubble to describe these MSAs is certainly justified from an ex post perspective: on average, these 20 cities experienced a decline in house prices from 2006 to 2010 of 46% compared to a decline of 12% for other cities.

The acceleration of the PLS market during the housing boom was an important driver of what happened in bubble MSAs. Figure 7 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 7. 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 bubble 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.

The standard for statistical proof of an asset price bubble is high (see, e.g., Fama (2014)). The Glaeser et al. (2008) model provides a potential test. As they argue, the presence of irrational home-buyers in areas where new construction is easy will predict a subesequent collapse in house prices and construction that will bring both below their pre-bubble levels. To test this prediction, columns 3 through 6 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, and so the relative level of construction and house prices between high and low NCL share MSAs in 2002 is an important benchmark.

As columns 4 through 6 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 and construction to fall even below their pre-boom steady-state levels. High NCL share MSAs saw prices and construction collapse from the 2006 peak, and the collapse was large enough to bring them below 2002 levels.[13] How could the acceleration of the PLS market have led to such large excesses in the most exposed areas? The next section shows that credit-fueled speculation was critical.

4 Credit and Asset Prices: the Speculation Mechanism

4.1 Volume

The acceleration of the PLS market in 2003 sparked a large increase in volume in high NCL share zip codes. Figure 8 presents coefficients from the estimation of equation 2 using log number of first-lien home purchase mortgage originations and log number of transactions as the left hand side variables.[14] 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 began exactly as the PLS market accelerated in 2003. A one standard deviation increase in the 2002 NCL share of a zip code was associated with a 5 to 7% increase in the number of purchase mortgages and volume in 2006.

Both the number of first-lien purchase mortgage originations and volume collapsed in 2007 in high NCL share zip codes. In 2009 and 2010, volume again rose in high NCL share zip codes, but the number of first-lien mortgages continued to collapse. This suggests that cash buyers moved into the market after the crash; we return to this point in Section 5.

Table 8 presents estimates of equation 3 using volume and first-lien mortgage originations

[13] An important caveat: the negative effect of defaults and foreclosures on house prices during the bust may bring prices even below long-run fundamental value in bubble cities. This force is not contained in the Glaeser et al. (2008) model.

[14] The analysis throughout this section uses only first-lien home purchase mortgages to capture a transaction financed with a mortgage. There was a large increase in second-lien "piggy-back" mortgages used to help purchase homes during this period, so the use of total purchase mortgages will significantly overstate the number of transactions financed with a mortgage.

as the left hand side variable. Specifications reported in columns 1 through 2 use as a left hand side variable 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 number of housing units as of 2000. Columns 3 and 4 use a similar left hand side variable with the change in first-lien mortgages in the numerator. As a percentage of the housing stock, there was a 1.5 to 2 percentage point increase in transaction volume for a one standard deviation increase in the 2002 NCL ratio. The increase in the number of first-lien purchase mortgages was almost identical. This implies that the relative rise in transaction volume in high NCL share zip codes was driven entirely by a rise in transactions using a first-lien purchase mortgage, highlighting the importance of the credit supply expansion in driving the rise in volume.

4.2 Speculators as marginal buyers

The mortgage-level TransUnion data set allows us to measure the marginal buyers brought into the housing market by the acceleration of the PLS market. To explore the marginal contribution of different groups, we utilize the following decomposition:

$$\frac{y_{z,m,0506} - y_{z,m,0102}}{y_{z,m,0102}} = \sum_i \frac{y^i_{z,m,0506} - y^i_{z,m,0102}}{y_{z,m,0102}}$$

The percentage change in any outcome variable y can be decomposed into the relative contribution from different groups i. The key outcome variable in this sub-section is the number of first-lien mortgages. For example, we can partition individuals by age to see what age group is driving the overall percentage change during the boom in the number of first-lien mortgages.

With this decomposition in mind, Table 9 presents estimates from the following specification for first-lien purchase mortgages:

$$\frac{firstlienmortgages^i_{z,m,0506} - firstlienmortgages^i_{z,m,0102}}{firstlienmortgages_{z,m,0102}} = \alpha_m + \beta^i NCL_{z,m,2002} + \varepsilon_{z,m} \quad (5)$$

where z is a zip code, m is an MSA, i is a partition of the population, and 0102 and 0506 represent the sum from 2001 to 2002 and 2005 to 2006, respectively.[15]

Columns 1 and 2 report specifications where i reflects the universe of first lien mortgages in the HMDA data and the TransUnion data, respectively. For the specifications reported in these columns, the left hand side variable is just the percentage change in first-lien mortgages in zip code z during the boom. A one standard deviation increase in the 2002 NCL share is associated with a 15% increase in first-lien mortgage originations, and the estimate is similar across the two data sets. This gives us confidence that the measurement of first-lien purchase mortgages in the TransUnion data set is similar to the HMDA data.[16]

The specifications reported in columns 3 through 5 focus on "flippers." Flippers are defined in two ways. First, a mortgage origination is classified as being taken out by a flipper if the individual taking out the mortgage in question also takes out another distinct first-lien purchase mortgage in a two year period around the origination in question. Second, a given first-lien purchase origination is classified as being taken out by a flipper if the mortgage is subsequently closed within a year, and there is no associated refinancing for the individual in the six months after the mortgage is closed. Our final definition of a flipper is the union of the two groups. As column 5 shows, flippers make up 11% of the total 14% increase. Flippers accounted for almost 80% of the relative increase in first-lien mortgage origination growth in high NCL share zip codes. Appendix Table 1 reports the specification without MSA fixed effects. In that specification, flippers make up 7% of the total 8% increase.

Figure 9 plots coefficients from a specification similar to equation 2 where the left hand side variable is first-lien purchase originations along with first-lien purchase originations by flippers. Both series are divided by total first-lien purchase originations as of 2002. As the figure shows, most of the relative increase in first-lien purchase mortgage originations in high

[15]The previous tables used 2000 to 2002 as the pre-period and 2004 to 2006 as the treatment period, but Table 9 switches to 2001 to 2002 and 2005 to 2006 because the TransUnion data are not available for the full year of 2000. All specifications reported in Table 9 include MSA fixed effects; the specifications without MSA fixed effects are reported in Appendix Tables 1 and 2.

[16]Appendix Figure 4 shows the entire time series of the NCL effect for TransUnion and HMDA data sets; the results are similar across the two data sets.

NCL share zip codes was due to flippers.

Table 10 shows that the marginal borrowers brought in by the PLS market were riskier and younger. Risk is measured in two ways: by the credit score as of 2000 and by whether the individual defaults on a mortgage in 2006 or afterward. As Panel A shows, most of the effect comes from individuals with a subprime credit score as of 2000 and individuals with no credit score as of 2000.[17] Individuals with a prime credit score as of 2000 in high NCL share zip codes did not witness stronger growth in first-lien mortgage originations relative to individuals with a prime credit score as of 2000 in low NCL share zip codes. Using a measure of risk based on ex post default, the results show that the entire rise in first-lien mortgage origination growth in high NCL share zip codes was driven by individuals who ex post defaulted on a mortgage.

As Panel B shows, there was a slight tilt toward young individuals in accounting for the relative increase in first-lien purchase mortgage origination growth in high NCL share zip codes. As shown in Table 1, the median age at origination of a first-lien mortgage from 2001 to 2010 was 40; the results in Panel B indicate that 65% of the relative growth in first-lien purchase originations in high NCL share zip codes was driven by individuals below 40 at origination.

Figure 10 presents the share of the total relative growth in first-lien mortgages in high NCL share zip codes during the boom for different sub-groups. The bars are the estimated β^i for each group i from equation 5 scaled by the total effect of 0.141 reported in column 2 of Table 9. As it shows, flippers accounted for 80% of the relative growth in high NCL share zip codes, and individuals with a subprime credit score as of 2000 accounted for 60%. Individuals who subsequently defaulted on mortgages accounted for all of the relative rise in first-lien mortgage origination growth in high NCL share zip codes.

[17] Column 4 of Panel A shows that 5.3% of the total 14.1% increase was driven by individuals that did not have a credit score as of 2000. When these individuals obtain a credit score in subsequent years, the average and median score are both in the near prime category.

4.3 Housing market optimism

Overall, the results in the sub-section above are consistent with the view that the expansion in credit supply allowed a small group of speculators to have large effects on the market. To put this in perspective, of all individuals in the TransUnion sample in 2005 and 2006, only 0.92% were flippers according to our broadest definition.

Evidence on housing market optimism from the Michigan survey supports this conclusion. As the left panel of Figure 11 shows, during the heart of the PLS acceleration period of 2003 to 2006, the fraction of individuals saying that now is a bad time to buy a house rose from 20% to almost 40%. Furthermore, this increase was driven in large part by individuals who became 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.[18]

However, consistent with Piazzesi and Schneider (2009), there was a smaller cluster of individuals who became more optimistic about housing because of price and credit considerations. This is shown in the right panel of Figure 11. Although the total fraction of individuals saying it is a favorable time to buy a home fell, the fraction saying it is a good time to buy because prices will rise or credit is easy increased.

How was housing market optimism related to the acceleration of the PLS market? The Michigan data allow for the use of cross-sectional variation in housing market optimism across MSAs by the 2002 NCL share to answer this question. Table 11 reports the following specification on the evolution of optimism on the housing market in high versus low NCL share MSAs during the housing boom:

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

[18]This does not dispute the observation in the literature that a change in expectations by *lenders* was an important part of the lending boom, see, e.g., Gerardi et al. (2008), Gennaioli et al. (2012), Landvoigt (2016), Kaplan et al. (2017). For example, an overall neglect of downside risks as in Gennaioli et al. (2012) could explain the rise in asset-backed securitization across many markets from 2003 to 2006.

where $\Delta Optimism_{m,BOOM}$ 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 1 and 3 present the OLS estimates, and columns 2 and 4 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 became 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 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 was driven by people who are pessimistic because of house price considerations. In Appendix Table 3, 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 was a relative increase in the fraction of individuals expressing pessimism in high house price growth MSAs during the boom.

The results are consistent with the view that 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.

5 The Crash

Speculators were instrumental in triggering the initial rise in defaults in 2006 and 2007. The left panel of Figure 12 plots the mortgage default rate for speculators in the highest quartile of the 2002 NCL share distribution, where speculators are defined as individuals who flipped houses in 2005 or 2006.[19] Default rates for flippers rose substantially from 2005 to 2006, and

[19]Quartiles in Figures 12 and 13 are weighted by 2006 total mortgage debt outstanding.

then reached almost 20% in 2007. By 2009, the default rate for flippers living in zip codes in the highest quartile of the 2002 NCL distribution was 35%. To put this into perspective, the aggregate default rate on mortgages peaked at 11% in 2009.[20]

The left panel of Figure 12 also plots the default rate for non-flippers living in zip codes in the top quartile of the 2002 NCL share distribution, and all individuals living in the bottom quartile. Default rates were significantly higher even among non-flippers most exposed to the acceleration of the PLS market in 2003. The significantly amplified housing boom and bust ended up affecting all mortgage holders in high NCL share zip codes markets.

A central idea from models of credit and belief heterogeneity is that the crash in asset prices can be particularly painful because more "optimistic" individuals can no longer participate in the market, and "pessimistic" individuals must step in to buy the asset before a price floor is reached. The right panel of Figure 12 shows evidence consistent with this idea. More specifically, it plots the share of total volume financed with a mortgage in the top and bottom quartile of the 2002 NCL distribution. The share was almost 100% during the housing boom in high NCL share zip codes, but then collapses to 75% from 2006 to 2009. This pattern suggests that the marginal buyer in high NCL share zip codes shifted from a leveraged optimist to a cash-buyer pessimist from the boom to the bust, which can help explain why prices dropped so spectacularly.

Figure 13 presents evidence that the overall U.S. mortgage default crisis was triggered by defaults in zip codes that were most prone to credit-induced speculation from the acceleration of the PLS market in 2003. From 2006 to 2007, total mortgage defaults rose substantially more in high NCL share zip codes relative to the rest of the country. The right panel shows the share of total defaults coming from the zip codes in the highest 2002 NCL share quartile, which rose by almost 5 percentage points from 2005 to 2007. By 2008 and 2009, the mortgage default crisis spread to the rest of the country (e.g., Ferreira and Gyourko (2015)), but Figure 13 shows that the earliest stage of the mortgage default crisis was driven by defaults in zip

[20] Bhutta et al. (2017) focus on non-prime mortgages originated in Arizona, California, Florida, and Nevada with 100% LTV ratios and find default rates close to 90%.

codes most exposed to the PLS market.

6 Discussion

6.1 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.[21] 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, speculation, and housing market optimism.[22]

Existing research uses plausibly exogenous variation in credit supply shocks to show that shifts in credit supply affect house prices.[23] In addition to using a distinct empirical strategy based on the PLS market, this study is the first to explore the speculation mechanism underlying the effect of a plausibly exogenous credit supply expansion on house prices.

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.[24] This literature shows that states that experienced the largest boom-bust cycle witnessed the largest increase in the participation of investors (Haughwout et al. (2014)), investors played an important role in explaining

[21]See, 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).

[22]The empirical strategy used here is closest to Nadauld and Sherlund (2013) who measure a zip code's exposure to the growth in securitization of mortgages by the five largest broker/dealer investment banks during the 2003 to 2005 period. They find that securitization affected mortgage originations and default rates, but they do not focus on house prices, construction, housing market optimism, or speculation.

[23]See, e.g., Adelino et al. (2014), Favara and Imbs (2015), and Di Maggio and Kermani (2017).

[24]See, e.g., Haughwout et al. (2014), Chinco and Mayer (2015), Bhutta (2015), DeFusco et al. (2018), and Gao et al. (2017).

the rise in household debt levels (Bhutta (2015)), short-term investors amplified volume and price movements in many markets (DeFusco et al. (2018)), and plausibly exogenous variation in speculative buying can explain the boom and bust in housing markets from 2000 to 2010 (Gao et al. (2017)). The main contribution of this study relative to this literature is to utilize a plausibly exogenous source of variation in exposure to the PLS market to show that credit supply expansion was instrumental in fueling short-term buying and speculation.[25]

There is a related body of research focusing on anomalous elastic housing supply MSAs with both a boom in construction and house prices.[26] 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 bubble MSAs such as Las Vegas and Phoenix.

Mian and Sufi (2009) use a within-county empirical strategy to show that zip codes with a higher share of subprime borrowers as of 1996 witnessed larger relative growth in mortgage originations for home purchase and house prices from 2002 to 2005. As shown in Table 3, within-MSAs, the 2002 NCL share of a zip code has a large and highly significant correlation with the share of subprime borrowers in 2000. This implies that many of the patterns shown in Mian and Sufi (2009) are related to the speculation channel emphasized here.

Recent research has focused on the distribution mortgage origination growth across the income distribution during the housing boom, and it shows that the share of total amount originated for home purchase increased for low income zip codes (e.g., Adelino et al. (2016), Mian and Sufi (2017a), Adelino et al. (2017)). For example, Figure 1 in Adelino et al. (2017) shows that zip codes in the bottom 40% of the 2002 income distribution increased their share of total purchase origination amounts from 22.7% to 27.3% from 2001 to 2006. To compare with this literature, Appendix Figure 5 splits zip codes into population-weighted quartiles

[25] Haughwout et al. (2014) show that borrowers during the housing boom with multiple first-liens were more likely to obtain credit from the non-prime part of the market.

[26] See, e.g., Glaeser et al. (2008), Davidoff (2013), Chinco and Mayer (2015), Nathanson and Zwick (2017), and Gao et al. (2017).

based on the 2002 NCL share, and it shows that zip codes in the top 2002 NCL share quartile saw an increase in the share of first-lien purchase mortgage originations from 25.1% to 28.1% from 2002 to 2006, and an increase in the share of total home purchase mortgage amounts from 20.6% to 25.4% from 2002 to 2006.

There is also a body of research exploring the rise of household debt across the income and credit score distribution during the housing boom (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 an outcome explored in this study. This study focuses on the extensive margin of house purchases, whereas the rise in the level of household debt from 2000 to 2007 was due primarily to homeowners borrowing against home equity (e.g., Mian and Sufi (2011), Mian and Sufi (2015), Bhutta and Keys (2016), Mian and Sufi (2017b)). The results presented here 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.

6.2 Conclusion

Theoretical models suggest that credit availability affects asset prices through a speculation channel. Belief heterogeneity, speculative behavior, and a large increase in volume are aspects of models in this literature. This study exploits variation across U.S. zip codes in exposure to the acceleration of the PLS market. It shows more exposed zip codes experienced a significant rise in house prices during the boom, and a spectacular crash during the bust. The credit boom was associated with a rise in volume and speculation, with belief heterogeneity playing an important role. Consistent with models of credit and speculation, the PLS market allowed a small group of individuals to have large effects on house prices from 2003 to 2010.

References

Abreu, Dilip and Markus K Brunnermeier, "Bubbles and crashes," *Econometrica*, 2003, *71* (1), 173–204.

Adelino, Manuel, Antoinette Schoar, and Felipe Severino, "Credit supply and house prices: Evidence from mortgage market segmentation," *Available at SSRN 1787252*, 2014.

_ , _ , and _ , "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.

Allen, Franklin and Douglas Gale, "Bubbles and crises," *The economic journal*, 2000, *110* (460), 236–255.

_ and Gary Gorton, "Churning bubbles," *The Review of Economic Studies*, 1993, *60* (4), 813–836.

Baron, Matthew and Wei Xiong, "Credit expansion and neglected crash risk," *The Quarterly Journal of Economics*, 2017, *132* (2), 713–764.

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.

_ and Benjamin J Keys, "Interest rates and equity extraction during the housing boom," *American Economic Review*, 2016, *106* (7), 1742–74.

_ and Benjamin Keys, "Eyes Wide Shut? The Moral Hazard of Mortgage Insurers During the Housing Boom," Technical Report, National Bureau of Economic Research 2018.

_ , Jane Dokko, and Hui Shan, "Consumer ruthlessness and mortgage default during the 2007 to 2009 housing bust," *The Journal of Finance*, 2017, *72* (6), 2433–2466.

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.

Cetorelli, Nicola and Stavros Peristiani, "The role of banks in asset securitization," *Federal Reserve Bank of New York Economic Policy Review*, 2012, *18* (2), 47–64.

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 2018.

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.

Fama, Eugene F, "Two pillars of asset pricing," *American Economic Review*, 2014, *104* (6), 1467–85.

Favara, Giovanni and Jean Imbs, "Credit supply and the price of housing," *The American Economic Review*, 2015, *105* (3), 958–992.

Ferreira, Fernando and Joseph Gyourko, "A new look at the US foreclosure crisis: Panel data evidence of prime and subprime borrowers from 1997 to 2012," Technical Report, National Bureau of Economic Research 2015.

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 2017.

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.

Gerardi, Kristopher, Andreas Lehnert, Shane M Sherlund, and Paul Willen, "Making sense of the subprime crisis," *Brookings Papers on Economic Activity*, 2008, *2008* (2), 69–159.

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, "Regulating the shadow banking system," *Brookings papers on economic activity*, 2010, *2010* (2), 261–297.

_ **and** _ , "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.

Harrison, J Michael and David M Kreps, "Speculative investor behavior in a stock market with heterogeneous expectations," *The Quarterly Journal of Economics*, 1978, *92* (2), 323–336.

Haughwout, Andrew, Donghoon Lee, Joseph S Tracy, and Wilbert Van der Klaauw, "Real estate investors, the leverage cycle, and the housing market crisis," 2014.

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

Kaplan, Greg, Kurt Mitman, and Giovanni L Violante, "The housing boom and bust: Model meets evidence," Technical Report, National Bureau of Economic Research 2017.

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.

Kindleberger, Charles P and Robert Z Aliber, *Manias, panics and crashes: A history of financial crises*, Palgrave Macmillan, 2005.

Landvoigt, Tim, "Financial intermediation, credit risk, and credit supply during the housing boom," *Working paper*, 2016.

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 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.

Miller, Edward M, "Risk, uncertainty, and divergence of opinion," *The Journal of finance*, 1977, *32* (4), 1151–1168.

Nadauld, Taylor D and Shane M Sherlund, "The impact of securitization on the expansion of subprime credit," *Journal of Financial Economics*, 2013, *107* (2), 454–476.

Nagel, Stefan, "The liquidity premium of near-money assets," *The Quarterly Journal of Economics*, 2016, *131* (4), 1927–1971.

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.

Pozsar, Zoltan, Tobias Adrian, Adam Ashcraft, and Hayley Boesky, "Shadow Banking," *Federal Reserve Bank of New York Economic Policy Review*, 2013, *19* (2), 1.

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

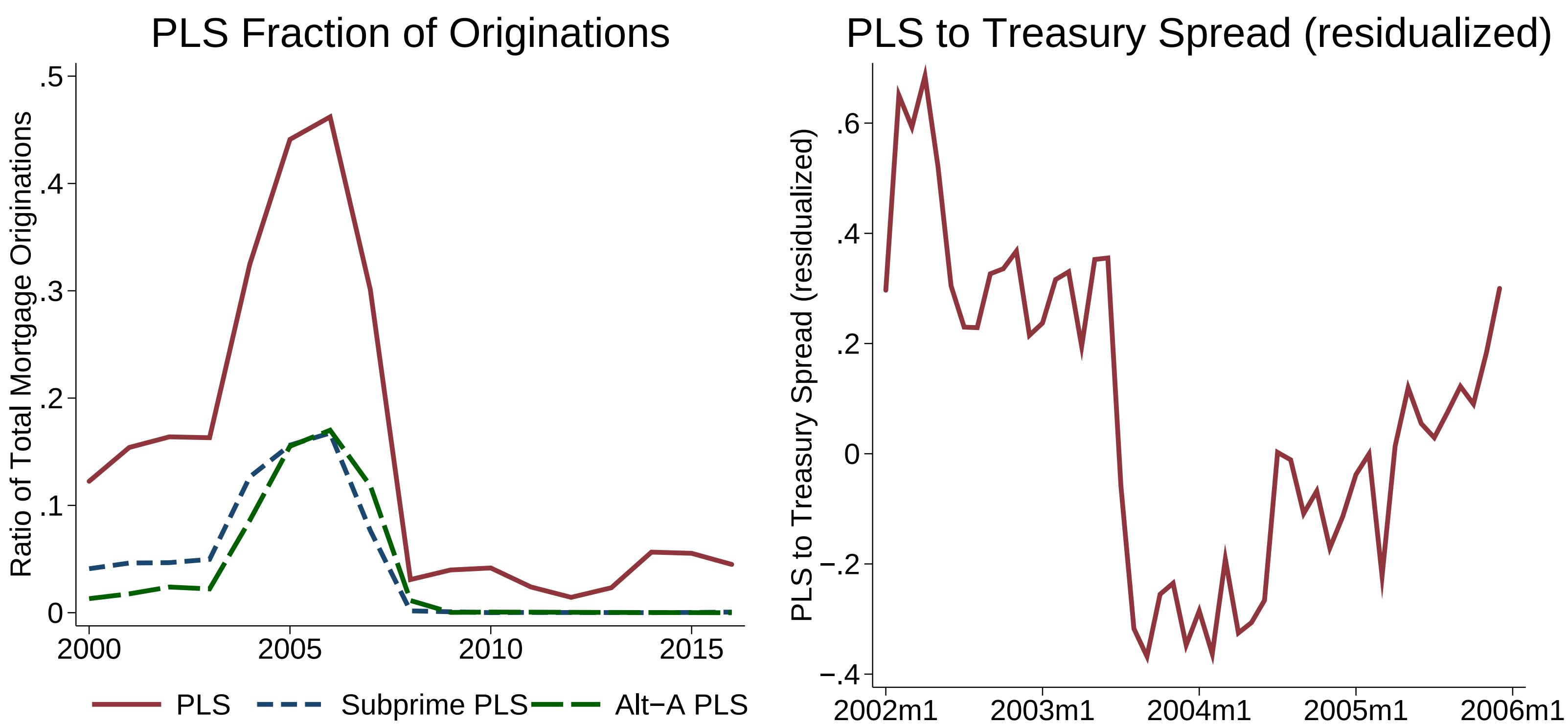
Saiz, Albert, "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

Scheinkman, Jose A and Wei Xiong, "Overconfidence and speculative bubbles," *Journal of political Economy*, 2003, *111* (6), 1183–1220.

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

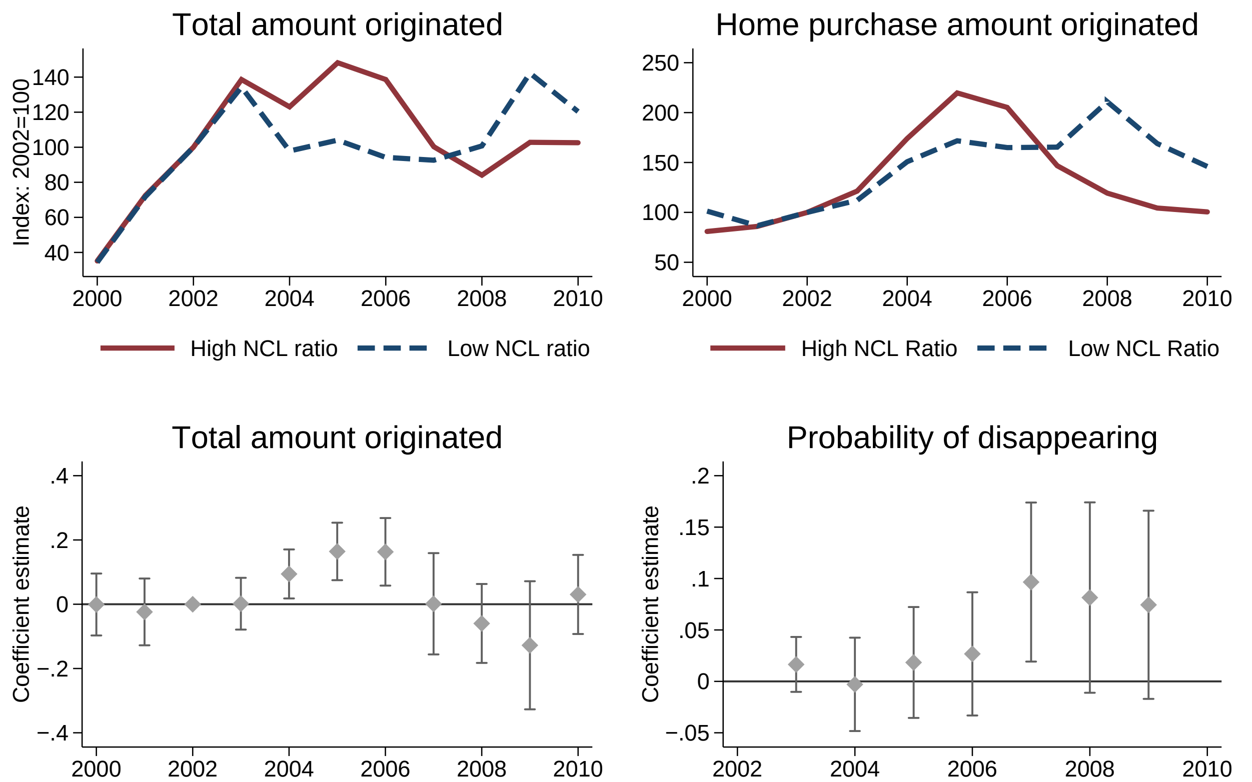
Xiao, Kairong, "Monetary Transmission through Shadow Banks," 2018.

Figure 1: Acceleration of Private Label Securitization of Mortgages



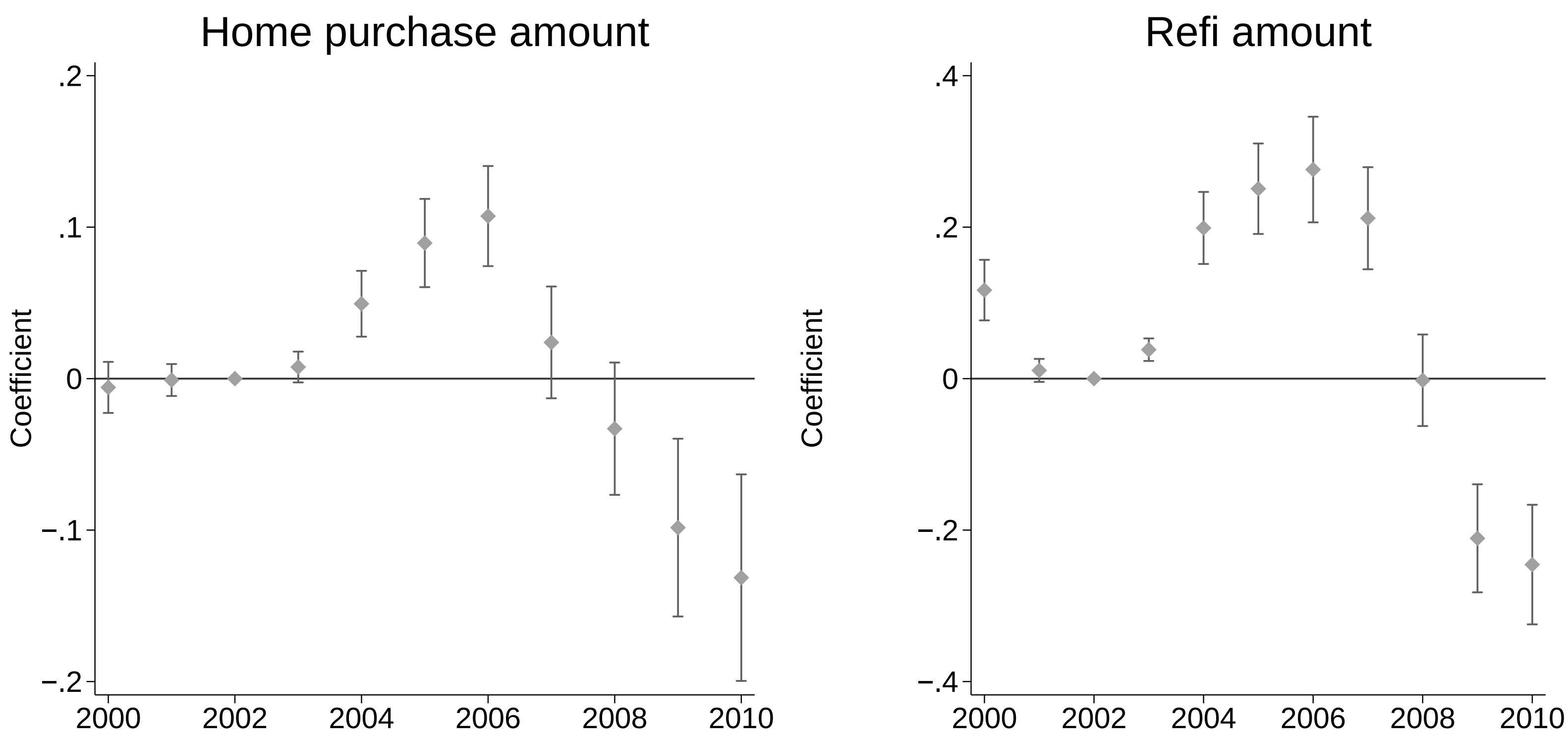
The left panel plots the share of total mortgage originations that were sold into private label securitization (PLS), subprime PLS, and Alt-A PLS. The right panel shows the average spread between mortgage interest rates in the private label securitization market and U.S. Treasuries from Justiniano et al. (2017), where characteristics of the mortgage are absorbed and the residual is plotted.

Figure 2: Relative Expansion of Mortgage Amounts Originated by High NCL Ratio Lenders



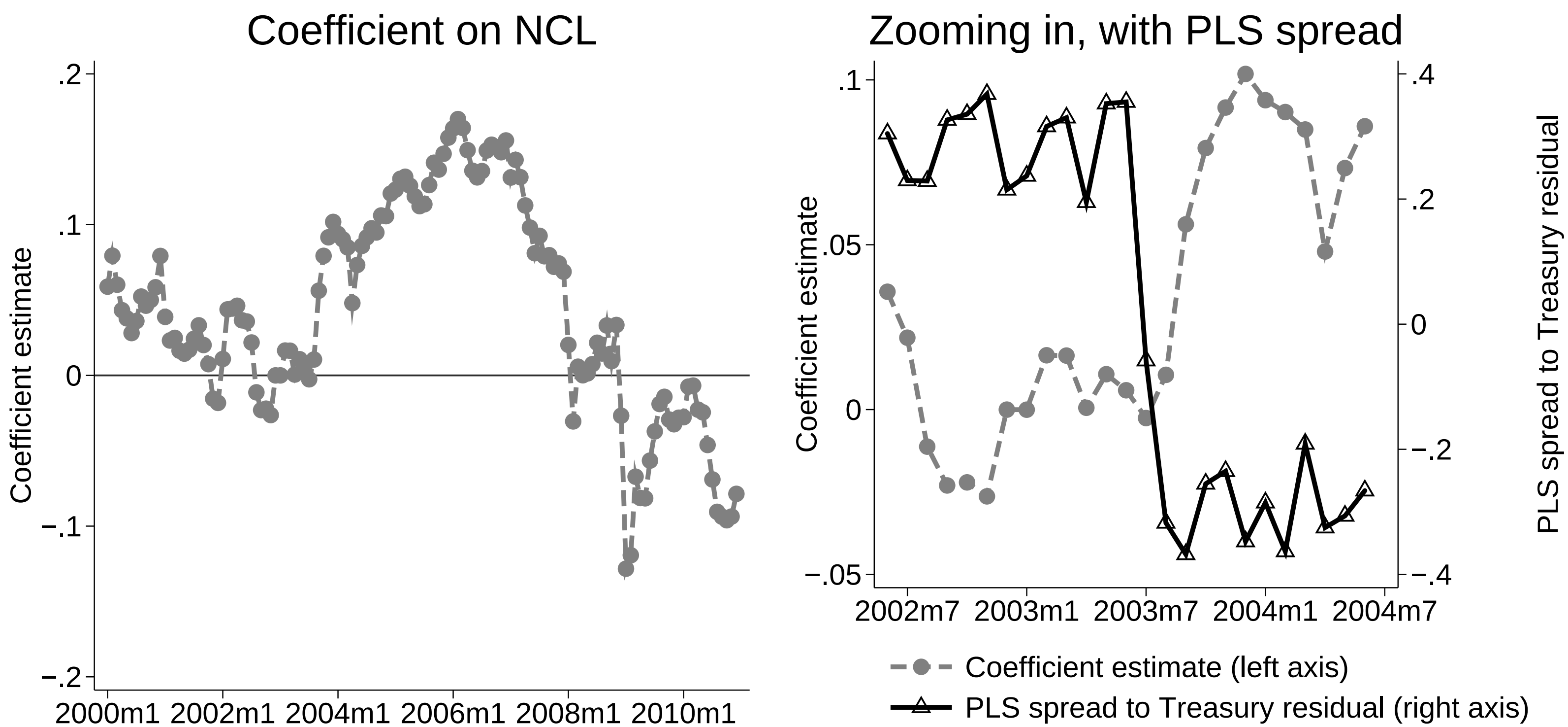
The top left panel plots total mortgage amount originated by lenders above and below the median non-core liabilities ratio (NCL) as of 2002. The top right panel plots home purchase mortgage amount originated by lenders above and below the median NCL as of 2002. The bottom left panel plots the coefficients $\{\beta_k\}$ 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 3: Zip-code Level Mortgage Amounts Originated by NCL Share: Panel Regressions



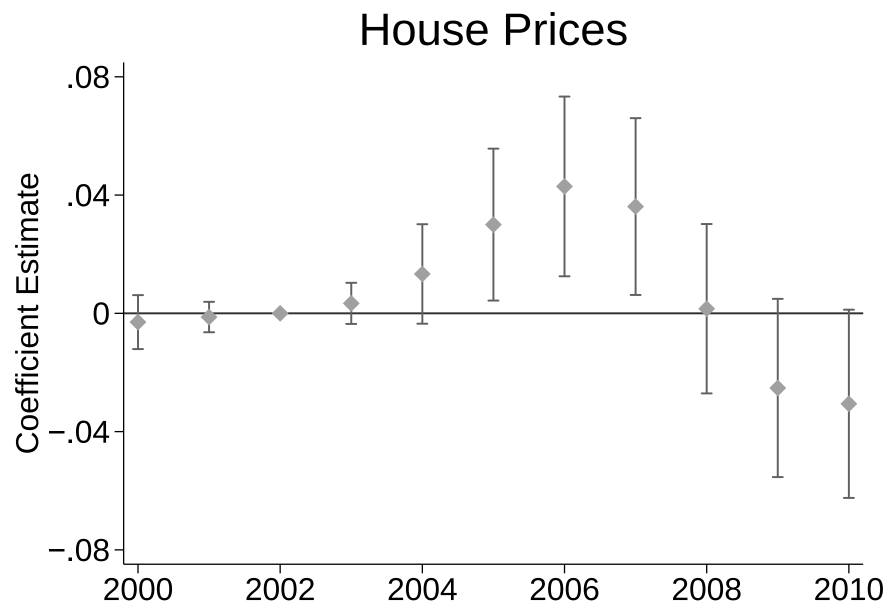
The panels plot the coefficients $\{\beta_k\}$ of the specification $ln(y_{z,t}) = \alpha_z + \gamma_t + \sum_{k\neq 2002} \mathbb{1}_{t=k}\beta_k NCL_{z,2002} + \varepsilon_{z,t}$ for zip code z in year t. $y_{z,t}$ is total mortgage amount originated in zip code z in year t. NCL at the zip code-level 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 4: MSA Monthly Level Mortgage Amounts Originated by NCL Share: Panel Regressions



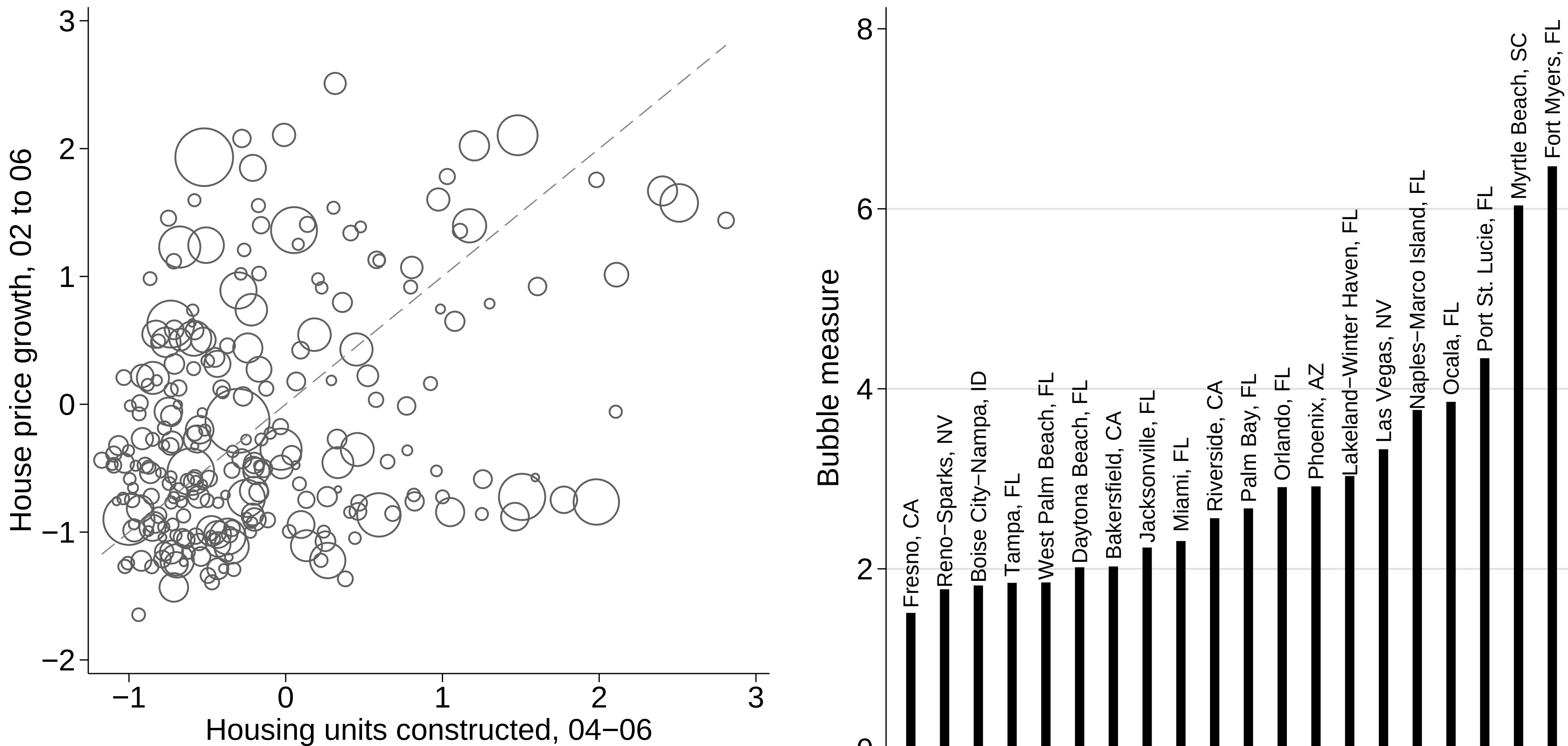
The left panel plots the coefficients $\{\beta_k\}$ of the specification $ln(y_{m,t}) = \alpha_m + \gamma_t + \sum_{k\neq 2002} \mathbb{1}_{t=k}\beta_k NCL_{m,2002} + \varepsilon_{m,t}$ for MSA m at time t. $y_{m,t}$ is total mortgage amount originated in MSA m in year t. The right panel zooms in around 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 5: House Price Growth by NCL share: Panel Regression



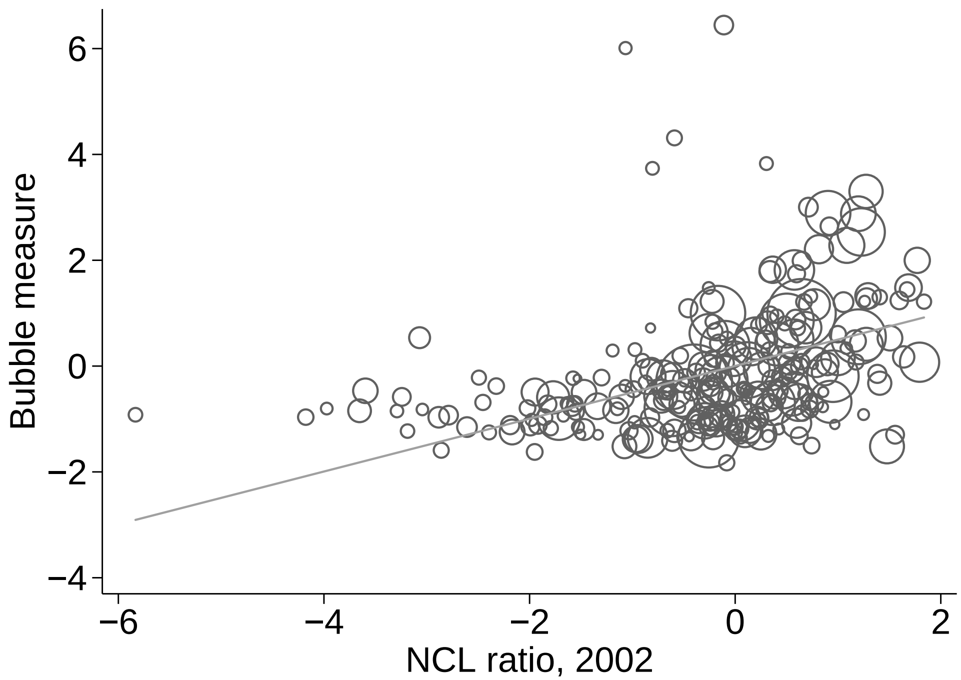
This figure plots the coefficients $\{\beta_k\}$ of the specification $ln(HP_{z,m,t}) = \alpha_{z,m} + \gamma_t + \sum_{k\neq 2002} \mathbb{1}_{t=k}\beta_k NCL_{z,m,2002} + \sum_{k\neq 2002} \mathbb{1}_{t=k}\beta_k SAIZ_{m,2002} + \varepsilon_{z,m,t}$ for zip code z, MSA m, and year t. $HP_{z,m,t}$ is house prices of zip code z and year t, $SAIZ_m$ is the Saiz (2010) elasticity measure of MSA m. 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 the zip code in 2000. 95% confidence intervals from robust standard errors are also plotted. Standard errors are clustered by MSA.

Figure 6: Bubble Cities: House Price Growth and Construction Activity



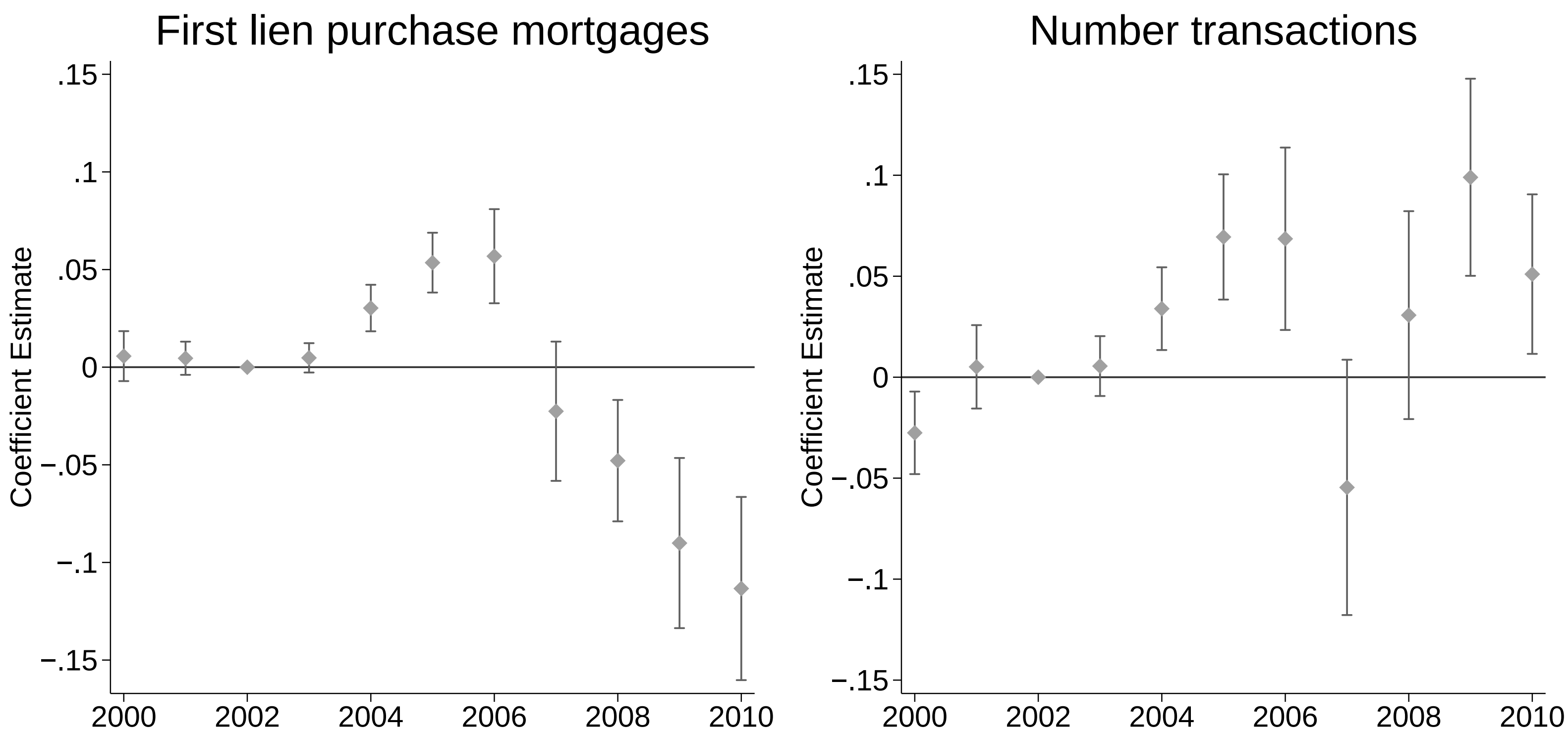
The left panel plots MSA-level house price growth from 2002 to 2006 against construction during the housing boom. The measure of construction is total units constructed from 2004 to 2006 scaled by the total number of units in the MSA as of 2000. Both measures are standardized to be mean zero and standard 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 rotating the left panel counter-clockwise 45 degrees, and then measuring the distance to the (0,0) point. See text for more details.

Figure 7: NCL Share and Bubble Cities



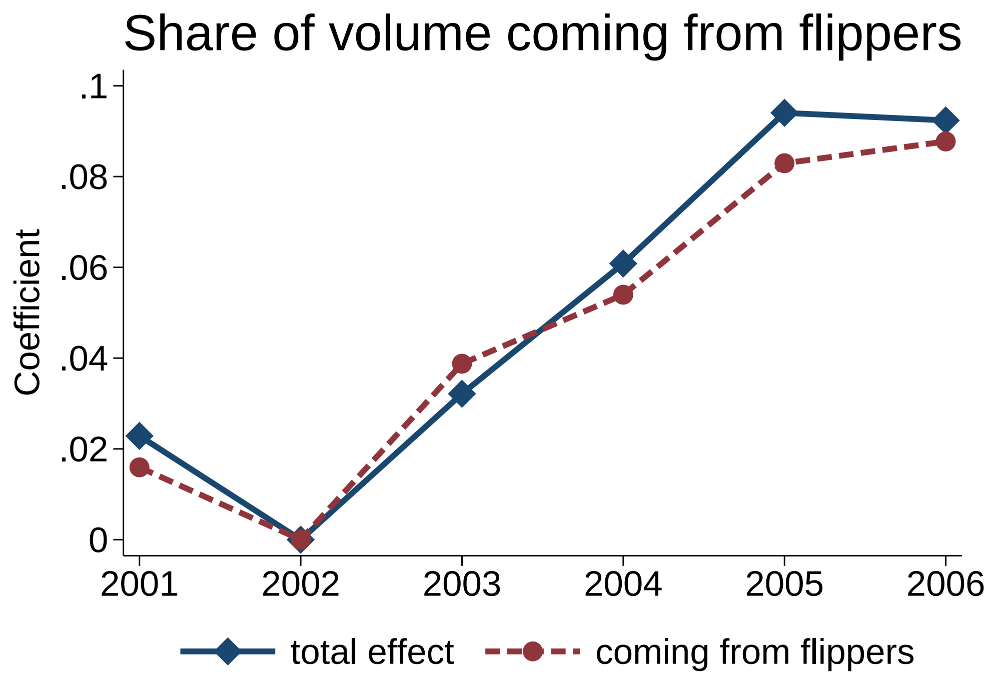
This figure plots for each MSA in the sample the bubble measure against the NCL ratio as of 2002. The bubble measure is constructed by starting with the scatter plot of house price growth from 2002 to 2006 against construction activity from 2004 to 2006, rotating the plot counter-clockwise 45 degrees, 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 8: Zip-code Transaction Volume by NCL Share: Panel Regressions



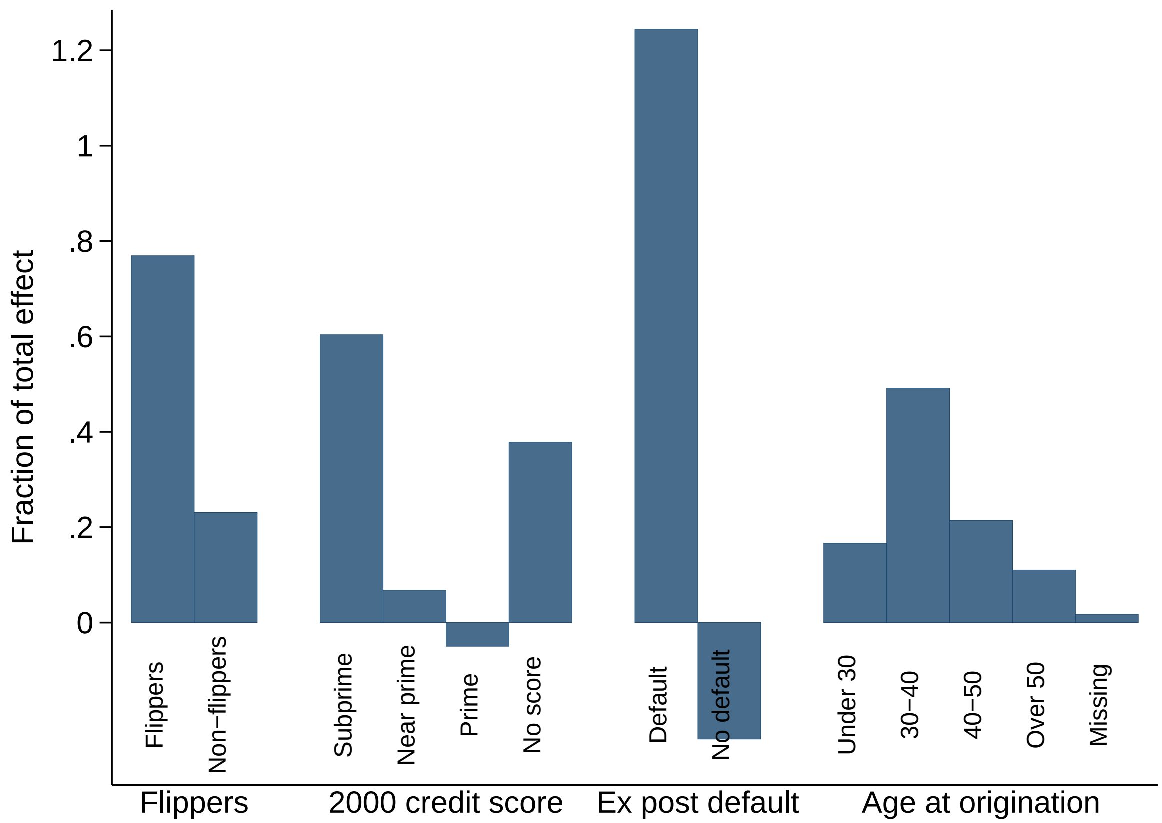
The panels plot the coefficients $\{\beta_k\}$ of the specification $ln(y_{z,t}) = \alpha_z + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{z,2002} + \varepsilon_{z,t}$ for zip code z in year t. $y_{z,t}$ in the left panel is the number of first lien mortgage originations for home purchase. $y_{z,t}$ 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 9: NCL Share and Flippers



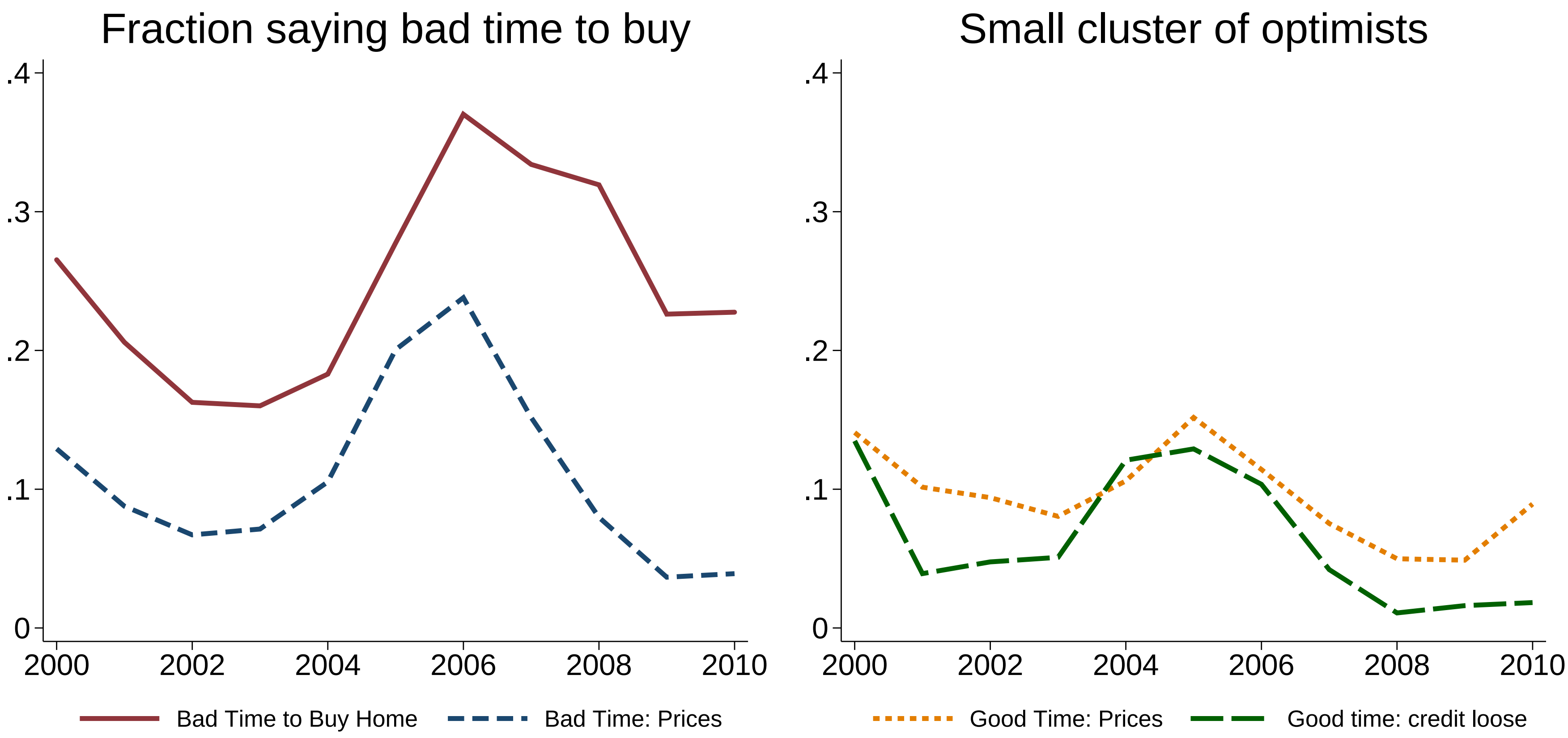
This figure plots the coefficients $\{\beta_k^i\}$ of the specification $yindex_{z,t}^i = \alpha_z + \gamma_t + \sum_{k\neq 2002} \mathbb{1}_{t=k}\beta_k^i NCL_{z,2002} + \varepsilon_{z,t}$ for zip code z in year t. $yindex_{z,t}^i$ is $\frac{firstlienmortgages_{z,t}^i - firstlienmortgages_{z,2002}^i}{firstlienmortgages_{z,2002}}$. For the blue solid line $yindex^i$ includes all first-lien mortgages in the TransUnion data. For the dotted red line, $yindex^i$ includes first-lien mortgages of flippers. Flippers are defined as individuals who take out at least 2 first-lien mortgages in a two year period, or take out a first-lien mortgage that is closed within one year with no associated refinancing. The regressions are weighted by the share of total occupied housing units in zip code z in 2000. Zip code level fixed effects included.

Figure 10: Decomposing Growth in First-lien Mortgages in High NCL Share Zip Codes



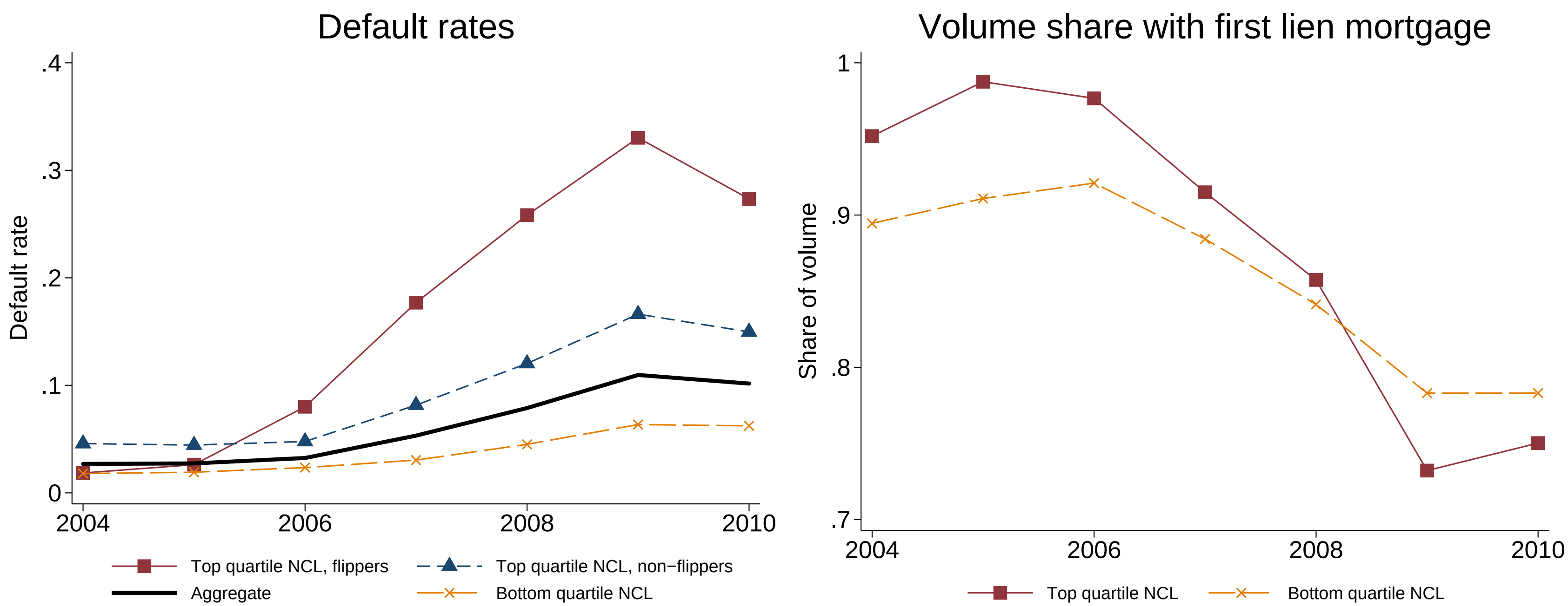
This figure plots the share of the relative growth in first-lien mortgage originations in high NCL share zip codes by certain groups. It is constructed by first estimating β^i from the following specification: $y^i_{z,m,BOOM} = \alpha_m + \beta^i NCL_{z,m,2002} + \varepsilon_{z,m}$ from zip code z and MSA m. The variable $y^i_{z,m,BOOM} = \frac{firstlienmortgages^i_{z,m,0506} - firstlienmortgages^i_{z,m,0102}}{firstlienmortgages_{z,m,0102}}$ where $firstlienmortgages^i$ is a subset of first-lien mortgages such as first lien mortgages taken out by flippers. For each group i, the coefficient β^i is divided by the total relative effect β estimated from: $\frac{firstlienmortgages_{z,m,0506} - firstlienmortgages_{z,m,0102}}{firstlienmortgages_{z,m,0102}} = \alpha_m + \beta NCL_{z,m,2002} + \varepsilon_{z,m}$.Plotted above is $\frac{\beta^i}{\beta}$, which is the share of the relative growth coming from group i.

Figure 11: Measures of Optimism on Housing Market from the Michigan Survey



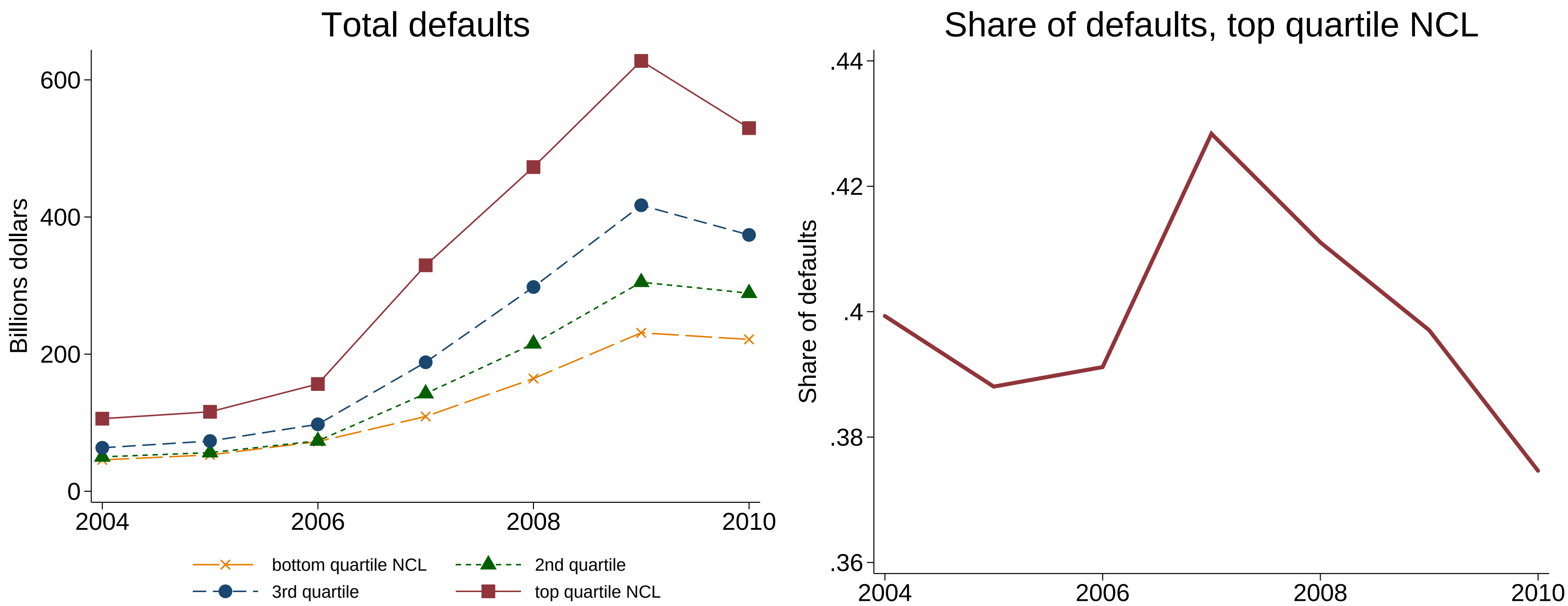
The left panel plots the share of individuals that respond to the question "Good or bad time to buy home" with the answers "bad time to buy a home" and "bad time to buy because of price considerations". The right panel plots the share of individuals saying it is "good time to buy because of price considerations" and "good time to buy because credit is loose." See text for more details.

Figure 12: NCL Share and the Crash



The left panel plots the mortgage default rate in zip codes by $NCL_{z,2002}$ quartiles. The quartiles are weighted by total mortgage debt outstanding as of 2006. Flippers are defined as individuals who take out at least 2 first-lien mortgages from 2005 to 2006, or take out a first-lien mortgage that is closed within one year with no associated refinancing from 2005 to 2006. 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 right panel plots the share of total volume with an associated first-lien mortgage recorded in HMDA.

Figure 13: Mortgage Defaults by NCL Share



The left panel plots delinquent mortgage debt in zip codes by $NCL_{z,2002}$ quartiles. The quartiles are weighted by total mortgage debt outstanding as of 2006. 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 right panel plots the share of delinquent mortgage debt in zip codes in the highest NCL exposure quartile.

Table 1: Summary Statistics

	Obs	Mean	SD	Median	P10	P90
Lender level						
2002 NCL ratio	5026	0.74	0.20	0.68	0.49	1.00
2002 Non-bank	5040	0.25	0.43	0.00	0.00	1.00
$\Delta_{02,05}$ PLS share	3950	0.15	0.26	0.08	-0.09	0.53
$\Delta_{02,05}$ ln (Amount originated)	3950	-0.02	0.73	-0.09	-0.46	0.62
Zip level						
2002 NCL Share	12427	0.77	0.05	0.77	0.71	0.82
Δ_{BOOM} (Home purchase amount originated)	12419	0.57	0.36	0.54	0.18	1.01
Δ_{BOOM} (Refinancing amount originated)	12400	0.32	0.53	0.23	-0.25	1.05
Δ_{BOOM} (First-lien mortgages, HMDA)	12418	0.14	0.28	0.12	-0.15	0.47
Δ_{BOOM} (Volume of housing transactions)	3727	0.16	0.29	0.12	-0.13	0.49
$\Delta_{02,06}$ (House Prices)	6619	0.37	0.22	0.36	0.10	0.67
Zip level: TransUnion data						
Δ_{BOOM} (First-lien mortgages, TransUnion)	9023	0.09	0.67	0.05	-0.69	0.92
Δ_{BOOM} (First-lien mortgages, HMDA)	9019	0.12	0.29	0.09	-0.21	0.47
Mortgage level: First-lien purchase						
Origination amount (thousands USD)	347905	187.86	176.31	149.20	40.55	360.00
VantageScore in 2000	295547	695.01	84.93	707.00	573.00	800.00
Age in year of origination	338304	42.18	12.68	40.00	27.00	59.00
Default in 2006 or after	347905	0.24	0.43	0.00	0.00	1.00
MSA Level						
2002 NCL Share	259	0.76	0.04	0.77	0.71	0.81
Housing Supply Elasticity	259	1.96	1.18	1.65	0.76	3.47
Constructed units, 04-06/Occupied units, 2000	259	0.03	0.03	0.02	0.01	0.08
Δ_{BOOM} Bad time to buy	259	0.07	0.13	0.06	-0.07	0.23
Δ_{BOOM} Bad time bc of prices	259	0.08	0.11	0.07	-0.04	0.21

This table presents summary statistics at the lender, zip, mortgage, and MSA level. $\Delta_{BOOM} Variable$ is defined as the log change in outcome y from 2000-2002 to 2004-2006. For the Zip level: TransUnion data panel, $\Delta_{BOOM} Variable$ is defined as the log change in outcome y from 2001-2002 to 2005-2006. This change is made because the TransUnion data are not available for 2000. Summary statistics at the lender level are weighted by total mortgage origination amount as of 2002. Summary statistics for the zip level and MSA level are weighted by total number of households as of 2000.

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)	(6)
					1998-2000	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	3947	3950	3947	3447	3433
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 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: 2002 NCL Share Correlations with Observable Variables

Covariates	Across MSA	Within MSA
2000 Deposits/Purchase amount originated	-1.11***	
	(.246)	
Saiz elasticity	-.262***	
	(.071)	
1998 NCL share	.849***	.842***
	(.024)	(.006)
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)
2000 Subprime share	.029***	.070***
	(.003)	(.001)

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 amount originated, 02 to 05		Bank-Zip-Code amount originated, 02 to 05	
	(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

	(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} Purch amount	Δ_{boom} Purch amount	Δ_{boom} Refi amount	Δ_{boom} Refi amount
2002 NCL Share	0.003	-0.005	0.086***	0.117***	0.224***	0.287***
	(0.012)	(0.009)	(0.016)	(0.017)	(0.030)	(0.030)
Level	MSA	MSA	Zip	Zip	Zip	Zip
MSA FE			N	Y	N	Y
N	338	338	12419	12419	12400	12400
R-sq	0.000	0.001	0.056	0.410	0.180	0.670

Standard errors in parentheses

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

Columns 1 and 2 present results for the specification $y_m = \alpha + \beta NCL_{m,2002} + \varepsilon_m$ for MSA m. 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). Columns 3 through 6 present 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 and refinancing mortgage amount originated. The NCL at the zip (MSA) 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 (MSA m). Regressions are weighted by the share of households in zip code z (MSA m) in year 2000. Standard errors are robust, and clustered at the MSA-level in columns 3 through 6.

Table 6: NCL Share and House Price Growth

	House Price Growth, 02 to 06					HP Growth, 06 to 10	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2002 NCL Share	0.059***	0.043**	0.151***	0.018*	0.063***	-0.084***	-0.049***
	(0.018)	(0.015)	(0.027)	(0.008)	(0.017)	(0.016)	(0.008)
Supply elasticity		-0.122***	0.791***				
		(0.017)	(0.151)				
2002 NCL Share X Supply elasticity			-0.055***		-0.027***		
			(0.010)		(0.007)		
MSA FE	N	N	N	Y	Y	N	Y
N	5540	5540	5540	5540	5540	5540	5540
R-sq	0.060	0.345	0.413	0.929	0.933	0.114	0.866

Standard errors in parentheses

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

Columns 1 through 5 present specifications of the form: $\Delta HP_{z,m,2002,2006} = \alpha_m + \beta NCL_{z,m,2002} + \beta_1 SAIZ_m + \beta_2 NCL_{z,m,2002} SAIZ_m + \varepsilon_{z,m}$ from zip code z and MSA m. $HP_{z,m,2002,2006}$ 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. Columns 6 and 7 present specifications of the form: $\Delta HP_{z,m,2006,2010} = \alpha_m + \beta NCL_{z,m,2002} + \varepsilon_{z,m}$ from zip code z and MSA m. $HP_{z,m,2006,2010}$ is the log change in house prices from 2006 to 2010 in zip code z. 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 7: NCL Share and Bubble MSAs

	(1) Bubble measure	(2) Bubble measure	(3) HP growth 02 to 10	(4) HP growth 02 to 10	(5) Δ units 09-11 minus 00-02	(6) Δ units 09-11 minus 00-02
2002 NCL share	0.439***	0.216**	-0.031*	-0.063***	-0.006***	-0.004***
	(0.078)	(0.069)	(0.012)	(0.013)	(0.001)	(0.001)
Housing supply elasticity	-0.283***	-0.198***	-0.017	-0.004	-0.000	0.001
	(0.057)	(0.044)	(0.011)	(0.010)	(0.001)	(0.001)
Census Division FE	N	Y	N	Y	N	Y
N	253	253	253	253	259	259
R-sq	0.290	0.513	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 anomaly measure, which is constructed by starting with the scatter plot of house price growth against construction growth, rotating the plot counter-clockwise 45 degrees, and then measuring the distance to the (0,0) point. See text for more details. y_m is the log change in house prices 2002 to 2010 in columns 3 and 4. 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 8: NCL Share and Change in Volume during Boom

	Δ_boom Volume per housing unit		Δ_boom First-lien per housing unit	
	(1)	(2)	(3)	(4)
2002 NCL Share	0.016***	0.016***	0.017***	0.018***
	(0.004)	(0.003)	(0.004)	(0.003)
MSA FE	N	Y	N	Y
N	3704	3704	3702	3702
R-sq	0.016	0.117	0.014	0.099

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 the sum of outcome y from 2004 to 2006 minus the sum of outcome y from 2000 to 2002, scaled by the number of housing units in the zip code as of 2000. The outcomes are transaction volume in columns 1 and 2 and the number of first lien mortgages originated in columns 3 and 4. 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 9: Who Are the Marginal Buyers? NCL Share and Flippers

	(1)	(2)	(3)	(4)	(5)
	Δ first-lien mortgages (HMDA)	Δ first-lien mortgages (TransUnion)	Δ first-lien mortgages, multiple houses	Δ first-lien mortgages, short-term	Δ first-lien mortgages, flippers
2002 NCL Share	0.157***	0.141***	0.067***	0.067***	0.108***
	(0.016)	(0.029)	(0.015)	(0.011)	(0.017)
MSA FE	Y	Y	Y	Y	Y
N	9020	9023	9023	9023	9023
R-sq	0.280	0.076	0.076	0.075	0.079

Regression results for specification: $y^i_{z,m,BOOM} = \alpha_m + \beta^i NCL_{z,m,2002} + \varepsilon_{z,m}$ from zip code z and MSA m. In columns 1 and 2, $y^i_{z,m,BOOM} = \frac{firstlienmortgages_{z,m,0506} - firstlienmortgages_{z,m,0102}}{firstlienmortgages_{z,m,0102}}$ for HMDA and Transunion data, respectively. For the rest of the columns, $y^i_{z,m,BOOM} = \frac{firstlienmortgages^i_{z,m,0506} - firstlienmortgages^i_{z,m,0102}}{firstlienmortgages_{z,m,0102}}$ where $firstlienmortgages^i$ is a subset of first lien mortgages such as first lien mortgages taken out by individuals who buy multiple homes or short-term investors. Individuals buying multiple homes in column 3 are defined to be individuals that obtain at least 2 first-lien purchase mortgages a two year period. Short-term buyers in column 4 are defined to be individuals taking out a first-lien purchase mortgages that is closed within a year after origination with no associated refinancing. In column 5, a flipper is defined to be someone that either buys multiple homes or is a short-term buyer. 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. All specifications include MSA fixed effects, and standard errors are clustered by MSA.

Table 10: Who Are the Marginal Buyers? By Measures of Risk and Age

Panel A: By 2000 credit score and ex post default

	(1)	(2)	(3)	(4)	(5)	(6)
	Δ first-lien mortgages, subprime	Δ first-lien mortgages, near prime	Δ first-lien mortgages, prime	Δ first-lien mortgages, no score	Δ first-lien mortgages, ex post default	Δ first-lien mortgages, ex post no default
2002 NCL Share	0.085***	0.010	-0.007	0.053***	0.175***	-0.034*
	(0.009)	(0.009)	(0.017)	(0.009)	(0.021)	(0.016)
MSA FE	Y	Y	Y	Y	Y	Y
N	9023	9023	9023	9023	9023	9023
R-sq	0.073	0.059	0.065	0.065	0.109	0.093

Panel B: by age at origination

	(1)	(2)	(3)	(4)	(5)
	Δ first-lien mortgages, lt 30	Δ first-lien mortgages, 30-40	Δ first-lien mortgages, 41-50	Δ first-lien mortgages, gt 50	Δ first-lien mortgages, missing
2002 NCL Share	0.023**	0.069***	0.030**	0.016	0.002
	(0.009)	(0.013)	(0.010)	(0.009)	(0.005)
MSA FE	Y	Y	Y	Y	Y
N	9023	9023	9023	9023	9023
R-sq	0.066	0.059	0.067	0.060	0.049

Standard errors in parentheses

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

Regression results for specification: $y^i_{z,m,BOOM} = \alpha_m + \beta^i NCL_{z,m,2002} + \varepsilon_{z,m}$ from zip code z and MSA m. $y^i_{z,m,BOOM} = \frac{firstlienmortgages^i_{z,m,0506} - firstlienmortgages^i_{z,m,0102}}{firstlienmortgages_{z,m,0102}}$ where $firstlienmortgages^i$ is a subset of first lien mortgages. In Panel A, first lien mortgages are split by credit score of the borrower as of 2000 and whether the individual defaults on a mortgage in 2006 or after. In Panel B, first lien mortgages are split by the age of the borrower as of 2001. 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. All specifications include MSA fixed effects, and standard errors are clustered by MSA.

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

	(1) Δ_{boom} Bad time to buy	(2) Δ_{boom} Bad time to buy	(3) Δ_{boom} Bad time to buy bc of prices	(4) Δ_{boom} Bad time to buy bc of prices
HP growth, 02 to 06	0.272***	0.336*	0.337***	0.287*
	(0.049)	(0.157)	(0.051)	(0.119)
Type	OLS	IV	OLS	IV
N	253	253	253	253
R-sq	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 $y_m = \alpha + \beta HPgrowth_{m,0206} + \varepsilon_m$ from MSA m. For columns 1 and 2, 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 3 and 4, 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. $HPgrowth_{m,0206}$ is house price growth in MSA m from 2002 to 2006. The regressions are weighted by the number of survey participants in an MSA m. In columns 2 and 4, $HPgrowth_{m,0206}$ is instrumented using the NCL share of the MSA as of 2002. 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. Standard errors are robust.