Preliminary and incomplete: Please do not circulate. Public draft March 2018

Credit Supply, Beliefs, and Speculation: Private Label Securitization and the Housing Cycle of 2000 to 2010

Atif Mian
Princeton University and NBER

Amir Sufi
University of Chicago Booth School of Business and NBER

February 2018

Abstract

Lenders that traditionally rely on non-core deposits to fund mortgages boosted mortgage originations substantially with the acceleration of the private label mortgage securitization market in 2003. Geographic areas with larger pre-2003 exposure to these lenders experienced a large relative increase in mortgage originations, house transaction volume, house prices, and construction activity from 2003 to 2006. Cities with large pre-2003 exposure to these lenders, such as Las Vegas and Phoenix, were significantly more likely to witness a simultaneous large increase in both house price growth and construction activity during the housing boom. The bust was particularly severe in these cities, bringing house prices and construction activity below pre-2003 levels. We are unable to find evidence that a rise in local optimism about house price growth triggered the rise in mortgage originations in these areas; on average, individuals living in these areas became more pessimistic about the housing market from 2003 to 2006. The results suggest that the acceleration of the private label market may have allowed a smaller group of optimistic buyers to have large effects on the housing market.

*This research was supported by funding from the Julis Rabinowitz Center For Public Policy and Finance at Princeton, and the Fama-Miller Center and Initiative on Global Markets at Chicago Booth. Michael Varley provided excellent research assistance. Mian: (609) 258 6718, atif@princeton.edu; Sufi: (773) 702 6148, amir.sufi@chicagobooth.edu

1 Introduction

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

This study presents evidence that the amplitude of the housing cycle is explained in part by the acceleration and subsequent collapse of the private-label mortgage backed securitization market (the PLS market hereafter).[1] Figure 1 shows the acceleration and subsequent collapse of the PLS market. Following the large mortgage refinancing wave of 2000 to 2002, the share of mortgages originated in the PLS market rose sharply. From 2003 to 2006, the PLS fraction of total mortgage originations rose from 16% of originations to 46%. As has been shown in the literature, the acceleration of the PLS market in 2003 was a credit supply-side phenomenon associated with a large increase in mortgage quantity and a significant decline in mortgage interest spreads (e.g., Levitin and Wachter (2013); Justiniano et al. (2017)).

What was the effect of the acceleration and collapse of the PLS market on the housing cycle? Our empirical strategy to address this question begins with cross-sectional variation across financial institutions in reliance on non-core deposits to fund loans. In particular, the acceleration of the PLS market after 2002 disproportionately reduced the cost of originating mortgages for financial institutions that traditionally relied on non-core deposit liability financing.[2] The extreme example of such financial institutions is the group of non-bank

[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.

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

mortgage lenders, such as Ameriquest Mortgage Company, that rely entirely on non-core deposit financing in the mortgage origination process.

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

The next step of the empirical strategy is to define geographic exposure across zip codes to high NCL lenders prior to the acceleration of the PLS market. In particular, the Home Mortgage Disclosure Act (HMDA) data is used to calculate the average NCL ratio for lenders originating mortgages in a given zip code, where the average NCL ratio is weighted by the amount of mortgage originations by the lender in the zip code in 2002. In this manner, we create a 2002 measure of cross-sectional variation across zip codes in exposure to the subsequent acceleration of the PLS market.

Zip codes more exposed to high NCL lenders see strong relative growth in mortgage originations from 2002 to 2006. This is true for both new home-purchase mortgages and for refinancing mortgages, and it is driven by a rise in originations sold into the private-label securitization market. The relative growth is especially strong for the extensive margin of home purchase. We use two measures of purchase intensity: the number of new purchase mortgages and a measure of volume from deed records collected by DataQuick. For both measures, the results show a significant increase in volume from 2002 to 2006 in zip codes

most exposed to high NCL lenders as of 2002. In terms of magnitudes, a one standard deviation increase in NCL exposure leads 15 to 25% increase in the growth of mortgage origination amounts and a 10 to 15% increase in transaction volume.

Our interpretation of these results is that the acceleration of the PLS market in 2003 represented a positive credit supply shock enabling buyers to obtain mortgage finance more easily (or cheaply) from 2003 to 2006, and this shock led to a large increase in transaction volume. We then explore how this shock affected house prices and residential construction. The results show that house prices and construction in high NCL areas experienced stronger relative growth from 2002 to 2006, and the timing matches closely the rise in mortgage origination growth. In terms of magnitudes, a one standard deviation increase in the NCL ratio leads to an increase in house price growth of 8 percentage points (12 percentage points in the most inelastic housing supply areas), and a rise in construction of 1.3% percent of the housing stock. The latter effect is almost one-half a standard deviation in the measure of construction.

One of the puzzling aspects of the housing boom from 2002 to 2006 is the presence of cities with elastic housing supply that simultaneously experienced large house price and construction growth (e.g., Glaeser et al. (2008), Davidoff (2013), Nathanson and Zwick (2017)). As shown Glaeser et al. (2008), these anomalous MSAs are difficult to explain in standard models with rational beliefs; if housing supply is elastic and there are a finite number of borrowers, then it is difficult to justify large increases in house prices.

The findings presented here suggest that the rise of the PLS market is a crucial factor in explaining this puzzle. We define "boom" cities as those that are in the top quartile of both the house price growth and construction growth distribution from 2002 to 2006. We show that exposure to NCL lenders as of 2002 predicts whether a city subsequently ends up as a boom city, and the magnitude is large. Further, cities with a large NCL share experience an "over-correction" to both house price growth and construction growth during the bust. By 2010, house prices and construction activity are below the 2002 level in high NCL share

cities. Glaeser et al. (2008) make exactly this prediction in a model with irrational beliefs: excessive building in elastic housing supply cities during the boom will lead to a steep decline in prices and construction when prices eventually fall. The results presented here suggest that credit provision fueled by the PLS market was a crucial ingredient in explaining what happened in these MSAs.[3]

We also investigate the evolution of housing market optimism in MSAs with a high NCL lender share. To conduct such an investigation, we use MSA-level measures of home buying optimism based on individual responses to the Michigan Survey of Consumers (Piazzesi and Schneider (2009)). We are unable to find evidence that individuals living in MSAs most exposed to high NCL lenders experienced a statistically significant relative change in attitudes toward home buying prior to the rise of the PLS market from 2000 to 2002. In other words, there is no evidence that a local "optimism shock" started the lending boom in high NCL lender share MSAs.

In fact, individuals living in high NCL MSAs on average become more pessimistic about the housing market from 2003 to 2006. On average, individuals increasingly report that it is a bad time to buy a house, and, in particular, they report that it is a bad time to buy a house because prices are too high. This evidence is inconsistent with the view that a general rise in optimism from 2003 to 2006 can explain relative mortgage origination, house price, and construction growth in high NCL share MSAs. Instead, the acceleration of the PLS market likely increased the purchasing power of a smaller group of more optimistic investors, many of whom may not have resided in the MSA itself (e.g., Chinco and Mayer (2015)). We are investigating this channel in ongoing work on this project.

Our interpretation of these results is that the rise of the PLS market had important effects on mortgage originations, transaction volume, prices, and construction activity. However,

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

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

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

Second, the timing of the relative growth in mortgage originations matches exactly the acceleration of the PLS market in mid-2003. There is no evidence of differential pre-trends in mortgage originations prior to 2003. Further, we use county-by-month HMDA data from the Federal Reserve to show that the relative increase in mortgage originations in high NCL share MSAs begins in September 2003, just as the PLS market accelerates in response to the decline in refinancing originations.[4] The relative growth in volume, house prices, and construction in high NCL areas also starts in 2003 with the acceleration of the PLS market. Further, the results are almost identical if we use a measure of exposure to high NCL lenders as of 1998 instead of 2002.[5] Third, we examine the main alternative hypothesis put forth

[4]This is closely related to the mortgage rate conundrum found by Justiniano et al. (2017). We discuss this in more detail below.

[5]These results are reported in the appendix.

in the literature: that a change in beliefs about house prices can explain the rise in house prices. As already mentioned, we are unable to find evidence that optimism about house price growth changed differentially in high NCL share MSAs before 2003.

There is a large body of research on the PLS market and the housing boom and bust from 2000 to 2010. We first present our main results, and then we discuss how our findings are related to this large body of research in Section 8.

2 Data and Summary Statistics

2.1 Data

The main data sets used in this study are at the lender, MSA, and zip code level. We begin with the Home Mortgage Disclosure Act (HMDA) data which records the universe of mortgage originations for mortgage originators that have an office within metropolitan statistical areas (MSAs).[6] 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.

More specifically, financial institutions report any mortgage loans to their regulatory agency, and they are given a unique ID number and agency code in the loan-level data. The financial institutions we focus on are regulated by the Office of the Comptroller of the Currency (OCC, agency code 1), the Federal Reserve System (FRS, agency code 2), and the Federal Deposit Insurance Corporation (FDIC, agency code 3); thrifts regulated by the Office of Thrift Supervision (OTS, agency code 4); and independent mortgage companies regulated by the Department of Housing and Urban Development (HUD, agency code 7). This only leaves out credit unions, who make up a small portion of lending, and institutions regulated by the Consumer Financial Protection Bureau, an agency that was only created

[6]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."

at the end of the decade as a response to the financial crisis. What we call banks are those institutions regulated by agencies corresponding to agency codes 1-4 since these are associated with a depository institution. Non-banks are non-depository independent mortgage lending companies and correspond to agency code 7.

We link lenders in the HMDA data to regulatory data filed by banks in the Report of Condition and Income (the Call Report) and by thrifts in the OTS Thrift Financial Report (TFR). Financial institutions that submit one of these forms are given a unique ID. If a bank (thrift) is part of a multi-bank (multi-thrift) holding company then each form provides a holding company ID which corresponds to the regulatory high holding company of the institution. For our analysis, we use the bank-holding or thift-holding company ID when an institution is part of a holding company and its unique ID otherwise. The HMDA ID is used for independent mortgage companies. Using a key of the HMDA Report IDs and the Call Report and TFR bank IDs provided to us by the Federal Reserve Board, we match the loan-level HMDA data to the bank level report data.[7]

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

The 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.

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

As a result, we do not have a measure of construction at the zip code level. Total volume comes from the DataQuick data base, subsequently purchased by CoreLogic. We follow closely the construction in DeFusco et al. (2017), and we include the same MSAs as in their sample. The data sets we use also include a measure of cash-out refinancing volume as reported from mortgage refinancing data from CoreLogic. The CoreLogic refinancing data is described in more detail in Mian and Sufi (2014).

2.2 Summary statistics

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

At the zip-code level, the 2002 NCL share is on average 0.78. The variables Δy_{BOOM} are constructed as follows. First, we add outcome y for zip code z in MSA m for years 2004 through 2006, and we then add outcome y for years 2000 through 2002. Δy_{BOOM} is defined to be the difference between the two. Housing supply elasticity comes from Saiz (2010) and is available at the MSA level. For the number of new housing permits, we measure total construction during the boom as total units constructed in an MSA from 2004 to 2006, scaled by total housing units in the MSA as of 2000.

3 Empirical Strategy

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

sources of variation.

3.1 Acceleration of the PLS market

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

The rise of the PLS mortgage market was part of the broader global pattern of the rise of securitization and shadow banking during the late 1990s and 2000s (e.g., Gorton and Metrick (2012); Gorton and Metrick (2013)). In fact, there is evidence that the rise of asset-backed securitization during this time period was broader than the housing market. In Appendix Figure 1, we show that originated amounts in non-mortgage-related asset-backed securities increased from less than $300 billion to almost $800 billion from 2002 to 2006. It is unlikely that specific views on the U.S. housing market can explain the rise in these alternative markets. We do not take a stand on the precise source of the aggregate credit supply shock driving the PLS market during this time-frame. Researchers have put forth a number of explanations including a global savings glut, a rise in income inequality, or lower uncertainty. We simply take it as given as an aggregate shock to the U.S. housing market.

The specific timing of when the PLS market accelerated is discussed in Justiniano et al. (2017). They show that the Federal Reserve signaled higher interest rates in the summer of 2003, which led to a collapse in mortgage refinancing for conforming GSE-backed mortgages. Justiniano et al. (2017) argue that mortgage originations for the PLS market accelerated directly after this episode, and mortgage interest spreads over Treasuries fell sharply. As we will show below, this is exactly the same time period we begin to see a relative expansion in

mortgage credit by lenders that rely on non-core deposits in their liability structure.

3.2 Lender-level exposure to the PLS market

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

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

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

as of 2002 were Ameriquest Mortgage Company, New Century Mortgage Corp, and American Home Mortgage Company.

3.3 Growth in mortgage originations: lender-level specifications

High NCL banks as of 2002 were more likely sell mortgages into the PLS market from 2002 to 2005. Figure 2 limits the sample to traditional banks (i.e., non-bank mortgage lenders are excluded) and plots the fraction of mortgages originated that are sold to a private institution. The outcome variable comes from the HMDA, which requires lenders to report to whom an originated loan is sold if it is sold within one year of origination. We follow Mian and Sufi (2009) and group together five categories that are a rough measure of mortgages sold into the PLS market.[8]

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

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

[8]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.

In columns 2 through 4, we present or main lender-level regression specifications relating mortgage origination growth 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 a 18% higher mortgage origination growth from 2002 to 2005.

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 loan 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 mortgage originations 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 mortgage originations concurrent with the acceleration of the PLS market.

Figure 3 presents evidence on mortgage originations from 2000 to 2010 based on the 2002 NCL ratio. The top two panels represent the averages of high and low NCL lender originations, where the two groups represent lenders above and below the median 2002 NCL ratio. As before, both the averages and the groups are formed using 2002 total mortgage originations as weights.

The top left panel uses total mortgage originations 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:

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

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

The evidence also suggests that there is a relative decline in mortgage lending by high NCL lenders when the PLS market collapses in 2007. However, this decline is underestimated because high NCL lenders are more likely to disappear from the sample after 2006. If a lender disappears, then it is not included in the sample for that year in the bottom left panel. The bottom right panel presents regression coefficients for a linear probability model that is similar to equation 1 except the left hand side variable is the probability of the lender being absent from the HMDA data in that year. As it shows, a one standard deviation increase in the 2002 NCL share implies a 10% higher probability of disappearing from the sample in 2007.[9]

3.4 Geographic exposure and MSA/zip code fixed effects

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

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

of mortgage originations in 2002.

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

These MSAs also tend to have less elastic housing supply with higher average house prices. MSAs with high NCL exposure tend to be higher income, although the coefficient is not statistically significant. However, within MSAs, high NCL lenders have the largest market share in lower income zip codes. The sign also 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 lower house price zip codes.

For the rest of the relationships, the across MSA and within MSA coefficients have the same sign. Areas with high NCL exposure tend to have lower homeownership rates, lower credit scores, and a younger population. The fact that high NCL lenders have large market share in zip codes with a lower fraction of individuals over the age of 65 is consistent with Becker (2007), who shows that seniors tend to save via deposits in local banks. Older zip codes are therefore "deposit-rich," and are less reliant on outside sources of funding.

All of these correlations make economic sense: deposit-poor areas are more likely to rely on mortgages originated by lenders that rely on external funding. In this sense, we do not want to control for these factors; they are the underlying source of variation in exposure to high NCL lenders.

However, this raises a concern: are shocks in these areas other than the acceleration of the PLS market responsible for the patterns we find? We examine this concern by focusing

on lender-MSA and lender-zip code level data sets. This allows us to include MSA or zip code fixed effects when estimating the effect of a high 2002 NCL share on origination 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 mortgage originations by lender b in geography g from 2002 to 2005. The geographical unit is an MSA in columns 1 and 2 of Table 4 and a zip code in columns 3 and 4.

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

Columns 3 and 4 conduct the same estimation at the zip code-lender level. The inclusion of zip code fixed effects boosts the R^2 by a factor of 12. The point estimate declines by 30%, but it remains economically large and statistically significant at the 0.001 level. If anything, this specification may "over-control" for the effect of the acceleration of the PLS market. If there is any spillover onto other lenders lending in the same zip code from the acceleration of the PLS market, the specification reported in column 4 will eliminate this effect. We view the estimate in column 4 as a particularly convincing result that the expansion of lending by high NCL lenders is not due to differential geographical exposure in their lending markets. The relative expansion of mortgage lending by high NCL share lenders appears to be due to the acceleration of the PLS market as opposed to local economic conditions.

4 Effects on Mortgage Origination Growth

Areas of the country with greater exposure to high NCL share lenders as of 2002 witnessed stronger mortgage origination growth and housing transaction volume from 2002 to 2005. In all of our subsequent analysis, the main right hand side variable is the average NCL share of mortgage lenders in the area as of 2002, where the average is weighted by the total amount of mortgage originations by the lender in the area in 2002. We refer to this variable as the 2002 NCL share of the area, and we construct it at both the zip code and MSA level depending on the outcome variable in question.

Figure 4 splits zip codes in the sample into high and low NCL share zip codes based on the population-weighted median, and it plots home purchase mortgage originations indexed to 2002. As it shows, there is no pre-trend from 2000 to 2002. From 2003 to 2005, high NCL exposure zip codes see stronger growth in mortgage originations for home purchase. From 2006 to 2010, mortgage originations in zip codes with high NCL exposure see a sharper decline.

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

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

The left hand side variable is the natural logarithm of the outcome measure in question in zip code lender z in year t. The coefficients β_k trace the relative growth of outcome y in zip codes with a high NCL share as of 2002. As in all specifications, $NCL_{z,2002}$ is normalized to have a standard deviation of one to ease interpretation.

The top left panel shows the coefficients using total mortgage originations as an outcome variable, and the top right panel uses only home purchase mortgage originations. Zip codes with high NCL exposure as of 2002 witness strong relative growth in both measures of mortgage originations from 2002 to 2006. Mortgage originations subsequently collapse after 2006, and by 2010 they are far below the 2002 level. Our interpretation of this pattern is

that the PLS market created a more amplified mortgage origination cycle from 2002 to 2010.

In Appendix Figure 3, we use an MSA by month level data set of mortgage originations from HMDA to examine more closely the timing of the increase in originations in high NCL exposure MSAs. These data are available from the Federal Reserve, and they are limited to the top 500 counties by mortgage originations in any year. As it shows, there is no pre-trend, and the relative expansion begins precisely in September 2003. As mentioned above, this is directly after the events of the summer of 2003 described in Justiniano et al. (2017). Long term interest rates began to rise after the Federal Reserve signaled an end to the easing cycle. The volume of refinancing originations of agency-backed mortgages declined and the PLS market rapidly accelerated. Mortgage interest spreads experienced an abrupt and large decline. The timing of the effect in Appendix Figure 3 supports the interpretation that the relative expansion in mortgage originations in high NCL exposure MSAs was due to the acceleration of the PLS market driven by supply-side factors.

The bottom two panels show that the relative rise in mortgage originations is boosting volume. We see a relative rise in volume using either the number of home purchase originations or transaction volume in the zip code as reported by DataQuick. Our interpretation of these results is that the rise of the PLS market led to a large boost in trading followed by a subsequent collapse in more exposed zip codes.

The left panel of Figure 6 shows results from a similar specification which uses as an outcome variable originations of mortgages that are refinancing existing mortgages. The right panel uses as an outcome variable the share of mortgages that are being refinanced with cash taken out (i.e., the principal balance of the mortgage increases with the home owner liquefying home equity). For both measures, we see a sharp rise in refinancing in zip codes with high NCL exposure from 2003 to 2006.

Tables 5 and 6 present regression coefficients from the following specification:

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

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 difference between the two. In other words, $\Delta y_{z,m,BOOM}$ is the 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 mortgage originations and volume 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 originations and volume. In the even columns, we add MSA fixed effects. The coefficient estimates on the 2002 NCL share variable become larger with the addition of MSA fixed effects. The MSA fixed effects are boosting the R^2 by at least a factor of four across the specifications, and their inclusion is actually boosting the estimated coefficient on the 2002 NCL share. This gives us confidence that MSA-level omitted shocks are not responsible for the effect of the NCL share on outcomes. In terms of magnitudes, the MSA fixed effects specification imply that a one standard deviation increase in the 2002 NCL share leads to a 25% increase in home purchase originations and an 8 to 15% increase in transactions depending on the measure used. Table 6 presents estimates from a similar specification using the measures of mortgage refinancing volume, and it shows similar results.

5 House Prices, Construction, and Boom MSAs

5.1 House prices

The acceleration of the PLS market at the end of 2003 led to a significant relative boost in mortgage originations and volume in high NCL share zip codes. What was the effect on house prices and construction activity? The left panel of Figure 7 presents estimates of β_k from the estimation of equation 2 with the logarithm of house prices as the left hand

side variable. As the estimates show, high NCL share zip codes experience positive relative growth in house prices in 2003, which then accelerates rapidly in 2004, 2005, and 2006. The estimates imply that a one standard deviation in the 2002 NCL share of a zip code leads to an 8 percentage point increase in house prices from 2002 to 2006.

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

Table 7 presents regression estimates from the following equation:

$$\Delta HP_{z,m,2002,2006} = \alpha_m + \beta NCL_{z,m,2002} + \gamma SAIZ_m + \delta NCL_{z,m,2002} SAIZ_m + \varepsilon_{z,m} \qquad (3)$$

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

As the estimate in column 3 shows, the 2002 NCL share in a zip code predicts high house price growth from 2002 to 2006. In terms of magnitudes, a one standard deviation increase in the 2002 NCL share leads to 6% higher house price growth. Column 4 includes the interaction between the 2002 NCL share and the measure of housing supply elasticity. As expected, the coefficient is negative: the effect of 2002 NCL share on house price growth

is weaker for zip codes located in more elastic housing supply MSAs.

The estimate on the 2002 NCL share variable in column 4 shows that the effect of the 2002 NCL share on house price growth is significantly stronger in the most inelastic housing supply MSAs. A one standard deviation increase in the 2002 NCL share boosts house price growth by 12% from 2002 to 2006 in the most inelastic housing supply MSAs. In MSAs where it is difficult to expand the supply of housing, the demand shock induced by the acceleration of the PLS market pushes house prices higher.

We conduct all of our regressions in this study in reduced form treating the 2002 NCL share as the underlying source of variation. We do so to avoid taking a specific stand on the structural relationship between the various outcomes. Of course, it is easy to calculate any relevant elasticity if one makes an assumption on the structural relationship. For example, if one assumes that the first stage is the growth in mortgage origination amounts on the 2002 NCL share instrument, and the second stage is the growth in house prices on the growth in mortgage origination amounts, then the elasticity of prices with respect to origination amounts is just the reduced form estimate divided by the first stage estimate. In this case, it would be $0.06/0.15 = 0.4$ (estimate in column 3 of Table 7 divided by estimate in column 1 of Table 5).

In columns 5 and 6, 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 county. Further, this effect is weaker in more elastic housing supply MSAs.

However, the absolute value of the coefficients drop between one-third to one-half in the specifications with MSA fixed effects. This is likely due to both statistical and economic reasons. The statistical reason can be seen with an examination of the increase in the R^2 when MSA fixed effects are included. For example, comparing columns 2 and 5, the R^2

[10]The **??** elasticity measure is defined only at the MSA level, and so the elasticity level drops out of the MSA fixed effects specification.

increases from 0.132 to 0.930. Such a large increase in the R^2 suggests that CoreLogic is smoothing its zip code level house price indices within MSAs.[11] 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 is the presence of spillovers. Zip codes within an MSA are not isolated islands. Price effects will therefore be muted as potential buyers search across neighboring zip codes. Such spillovers suggest that the use of within-MSA variation may lead to an underestimate of the effect of the NCL share on house price growth.

5.2 Construction

The right panel of Figure 7 examines construction activity as measured by new building permits collected by the Census. Data on new building permits are available only at the county level, and so we estimate equation 2 at the MSA level instead of the zip code level.

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

Table 8 presents estimates from a regression specification similar to equation 3 above with two differences. First, the specifications are estimated at the MSA level instead of the zip code level. Second, the left hand side variable is cumulative number of housing units constructed in 2004, 2005, and 2006 scaled by the number of housing units in the MSA in 2002. The estimate reported in column 2 implies that a one standard deviation increase in the 2002 NCL share of an MSA is associated with a 1.2 percentage point increase in the

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

number of units constructed from 2004 to 2006 as a fraction of the 2000 housing stock. This is almost one-half a standard deviation of the left hand side variable. As column 3 shows, the coefficient estimate is almost identical when including housing supply elasticity of the MSA as a control variable.

In column 4, we interact the 2002 NCL share with housing supply elasticity. The coefficient estimate on the interaction is not statistically different than zero. This is a puzzle. We would expect a stronger housing demand shock induced by the acceleration of the PLS market to induce more construction in MSAs with elastic housing supply. The sign of the coefficient is consistent with this prediction, but the statistical confidence and economic magnitude are low. We believe this result is related to findings in Glaeser et al. (2008), Davidoff (2013), and Nathanson and Zwick (2017) who find evidence of anomalous MSAs that experienced both high house price growth and construction growth from 2002 to 2006. We explore this issue in more detail in the next sub-section.

5.3 The anomalous boom MSAs

Motivated by Glaeser et al. (2008), Davidoff (2013), and Nathanson and Zwick (2017), we define a boom MSA as one that is in the top population-weighted quartile of both the house price growth and construction growth distribution during the housing boom.[12] There are a total of 30 MSAs in the boom category; Phoenix, Las Vegas, Orlando, and the central valley and inland empire California MSAs of Merced, Modesto, San Bernadino, Stockton, and Bakersfield are in this group. These MSAs experienced a simultaneous large increase in construction and house price growth from 2002 to 2006.

Figure 8 shows the relationship of house price growth (left panel) and construction activity (right panel) with housing supply elasticity, where boom MSAs are given a diamond marker in the scatter plot. As the left panel shows, while some of the boom cities have inelastic housing supply, others are closer to the middle of the housing supply elasticity

[12]More specifically, a boom MSA is in the top population-weighted quartile of both left hand side variables in Tables 7 and 8.

distribution (Nathanson and Zwick (2017)). All of the boom cities have higher house price growth than their housing supply elasticity would predict. The right panel shows that construction activity was not substantially higher in more elastic housing supply MSAs, a point raised by Davidoff (2013) and Nathanson and Zwick (2017). The lack of a positive relationship between construction activity and housing supply elasticity is due in part to boom MSAs, which had the largest increases in construction despite being in the middle of the housing supply elasticity distribution.

Glaeser et al. (2008) build a model to understand the existence of MSAs with a simultaneous increase in house prices and construction. As they put it, "rational bubbles can exist when the supply of housing is fixed, but not with elastic supply and a finite number of potential home buyers." They examine a model with an irrational temporary increase in optimism about future prices, and they show that such a shock can help explain why prices and construction may increase rapidly in MSAs with elastic housing supply. They also make a prediction that the crash will be particularly severe in these MSAs, because there will be overbuilding during the bubble.

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

The left panel of Figure 9 uses MSA-level data to plot the correlation between house price growth from 2002 to 2006 and the 2002 NCL share as of 2002. There is a strong positive correlation, consistent with the results in Table 7. In addition, we separately identify boom MSAs using diamond markers in the scatter plot. As the left panel makes clear, boom MSAs have on average a very high mean NCL share as of 2002. In fact, only one of the 30 boom MSAs has an 2002 NCL share below the unweighted sample average. The right panel shows the analogous plot for construction. As in Table 8, the correlation is positive. And as in the left panel, the boom MSAs play an important role in driving this correlation.

In Table 9, we estimate a linear probability model where we directly relate whether an MSA becomes a boom MSA to the NCL share as of 2002.

$$Prob(Boom_m = 1) = \alpha + \beta_0 SAIZ_m + \beta_1 NCL_{m,2002} + \beta_2 NCL_{m,2002} SAIZ_m + \varepsilon_m$$

where $Boom_m$ is equal to one if an MSA ends up in the top quartile of both the house price growth and construction growth distribution. As column 2 shows, a one standard deviation increase in the NCL share as of 2002 leads to a 10 percentage point increase in the probability of becoming a boom MSA. This is a remarkably large effect given that the mean of the left hand side variable is 11%. Controlling for housing supply elasticity has almost no effect on this estimate.

These results suggest that the puzzle of boom MSAs is directly related to the presence of high NCL lenders in these areas, and their resulting expansion in mortgage lending from 2002 to 2006. To the best of our knowledge, these results are the first in the literature to connect what happened in these MSAs to the PLS market.

We also test an additional prediction from the Glaeser et al. (2008) model: the collapse in house prices in boom MSAs may bring house prices even below their initial level before the boom because of overbuilding. Columns 4 through 6 regress house price growth from 2002 to 2010 on the NCL share as of 2002 to test this prediction. The underlying assumption is that the housing market was in a steady state equilibrium in 2002 before the acceleration of the PLS market. If housing supply were completely inelastic, an unexpected demand shock and reversal would lead to a boom and bust in house prices, but the long run level would remain constant at the 2002 level.

As columns 5 and 6 show, house price growth from 2002 to 2010 was relatively lower in MSAs with a high 2002 NCL share. The results are less precise, and are only statistically distinct from zero at the 10% level of confidence. In relative terms, high 2002 NCL share MSAs witnessed house prices fall during the bust below their 2002 level. These results

suggest that the acceleration of the PLS market from 2002 to 2006 led to housing market excesses in high NCL share MSAs that eventually caused house prices to fall even below their pre-boom house price level.

6 Housing Market Optimism

Some research points to optimism about house prices as a potential source of the housing cycle from 2000 to 2010. How is housing market optimism related to the acceleration of the PLS market in 2003? This section explores this question.

6.1 Data and aggregate patterns

Large sample data sets on beliefs about the evolution of house prices and the housing market prior to 2007 are unavailable.[13] Here, we follow Piazzesi and Schneider (2009) and use the Michigan Survey of Consumers to measure beliefs about the housing market. In particular, the Michigan survey asks the following question to respondents: "Generally speaking, do you think now is a good time or bad time to buy a house?" Almost 99% of survey respondents answer this question by saying it is either a "good" or "bad" time. In addition to this question, respondents are asked a follow up question: "Why do you say so?" In response to this question, survey questioners record a number of different responses, which we classify into sub-groups as follows.

If an individual responds that now is a good time to buy a house because "prices are going up," "prices are low/stable/not too high," "prices won't get any lower," or "good buys available," we classify the response as being favorable because of house price considerations. If an individual answers that now is a bad time to buy a house because "prices will fall later, will come down" or "prices are too high, houses cost too much," we classify the response as being unfavorable because of house price considerations. Note that there are many other

[13]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 **?** and covers only four metropolitan areas.

responses to the second question; the most common response given is that now is a good time to buy a home because "interest rates are low." As a result, the fractions of respondents saying now is a good time versus a bad time to buy because of house prices considerations will not add up to one. Further, the total fraction mentioning prices as a consideration will vary over time.

Figure 10 shows that evolution of answers to the Michigan questions from 2000 to 2010 for the full sample. During the PLS acceleration period of 2003 to 2006, the fraction of individuals saying that now is a bad time to buy a house rises from 20% to almost 40%. This result matches the finding in Piazzesi and Schneider (2009). Further, this increase is driven in large part by individuals who are sour on the housing market because of high prices. These findings match those of Piazzesi and Schneider (2009); they are inconsistent with the view that general optimism about housing was responsible for the rise in house prices from 2003 to 2006.

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

6.2 MSA-level analysis

How do the aggregate patterns on housing market optimism relate to the acceleration of the PLS market? We focus on cross-sectional variation across MSAs in the NCL share to explore this question. To the best of our knowledge, this study is the first to examine cross-sectional variation across MSAs in the answers to home buying questions in the Michigan Survey of Consumers.[14] In Figure 11, we report the coefficient estimates of β_k when estimating

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

equations similar to equation 1 but using as a left hand side variable the average response to the Michigan survey question at the MSA-by-year level. The goal of this exercise is to see whether MSAs with a higher NCL share as of 2002 see any differences in home buying attitudes during the sample period.

The left hand side variable in the top left panel is the fraction of individuals in the MSA saying that now is a bad time to buy a home. There is no statistically significant pre-trend. During the boom, individuals living in high 2002 NCL share MSAs on average increasingly report that now is a bad time to buy a home. The top right panel shows that this effect is driven largely by individuals saying it is a bad time to buy a home because of price considerations. Among individuals living in high 2002 NCL share MSAs, we see no evidence of a general rise in optimism concerning the housing market. In fact, we see the opposite. Recall that these same MSAs are experiencing a rapid rise in mortgage originations, house prices, and construction activity.

The bottom two panels test whether there is a smaller cluster of individuals living in these MSAs that become more optimistic because of price considerations. There is some suggesting evidence, but it is weak. The fraction of individuals saying it is a good time to buy because of price considerations shows no pre-trend, and there does not appear to be significant movement during the boom. If we examine the fraction mentioning prices conditional on saying it is a good time to buy a home, we find some evidence in 2005 of a rise in optimism.

Table 10 reports specifications on the evolution of optimism on the housing market in high versus low NCL share MSAs both before and during the housing boom. The first column shows that from 2000 to 2002, there is no statistically significant relative change in the fraction of individuals in an MSA with a high NCL share saying that it is a good time to buy a home because of price considerations. The second column shows there is no relative change in the index of consumer expectations in these areas; the ICE is a common measure from the Michigan survey used to measure expectations of economic conditions. We

are unable to find evidence that a local optimism shock explains the rise of the mortgage originatons, house prices, or construction in high NCL share MSAs.

Columns 3 through 6 present estimates of the following specification:

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

where $\Delta MichMeasure_{m,BOOM}$ is the MSA-level average of the survey responses to a given Michigan question in MSA m in years 2004 through 2006 minus the average of the survey responses to the same question in MSA m in years 2000 to 2003. The results in columns 3 through 6 show evidence consistent with Figure 11; high NCL share MSAs witness a rise in the fraction of individuals saying that now is a bad time to buy a house, and this is driven by individuals who say it is a bad time to buy a house because of price considerations. The results show no statistically robust increase in the fraction of individuals citing optimism on house prices in high NCL share MSAs.

These results are inconsistent with the view that a general rise in optimism on the housing market in high NCL share MSAs instigated the housing boom in these cities. We believe the more likely explanation is that the acceleration of the PLS market transferred purchasing power to a smaller group of optimistic home buyers, some of whom may not reside in the MSA. The evidence in Chinco and Mayer (2015) suggests this is an important channel. They show evidence that house prices in boom MSAs were driven at least in part by investors from outside the MSA. One potential explanation of our findings is that the acceleration of the PLS market in high NCL share MSAs allowed out-of-town optimistic buyers to purchase homes while in-town residents were increasingly selling or sitting out. We are exploring this channel in ongoing work on this project.

7 Delinquencies

As shown above, there was a collapse in the housing market in high NCL share zip codes from 2006 to 2010. Perhaps not surprisingly, there was also a large relative increase in mortgage default rates. The top left and right panel of Figure 12 presents coefficient estimates of β_k when estimating an equation similar to equation 3 but with the household default rate on debt as the left hand side variable.

During the boom, default rates in high NCL share zipcodes actually fall, consistent with a housing boom allowing individuals to avoid default through refinancing (e.g., Mian and Sufi (2011)). However, default rates then rise sharply from 2006 to 2009. In 2009, a one standard deviation increase in the NCL share of a zip code as of 2002 predicts a three percentage point increase in the default rate. The results are almost identical with and without MSA-by-year fixed effects. In other words, regional patterns are not responsible for the effect of the PLS market expansion on subsequent default rates.

In the bottom two panels of Figure 12, we show that the top population-weighted quartile of zip codes by the 2002 NCL share accounted for a substantial fraction of total delinquencies during the Great Recession. As the bottom right panel shows, the share of total delinquencies accounted for by the top quartile of zip codes based on the NCL share increased from under 30% to over 40% from 2005 to 2008.

One way to benchmark this result is to calculate the share of total delinquencies coming directly from the PLS market. One can obtain such an estimate using data from Foote et al. (2016) and Standard and Poor's. As of 2007, Foote et al. (2016) estimate that Alt-A and subprime mortgage originations accounted for approximately 17% of the stock of outstanding debt. Standard and Poor's estimate that delinquency rates on non-agency mortgages was approximately 40% for the 2006 vintage relative to 15% for agency mortgage debt. These numbers imply that the PLS market accounted for 35% of all delinquencies. Of course, some of the delinquencies on agency debt could have been the result of the acceleration of the PLS market through indirect channels. For example, if the acceleration of the PLS market led to

a rise and subsequent collapse in house prices, then even individuals with agency mortgages may default. Our interpretation of these results is that the PLS market was an important driver of mortgage defaults, but it is not the complete story.

8 Discussion

8.1 Relation to existing research

The findings presented here are related to a large body of research on the PLS market and subprime mortgages in particular (e.g., Keys et al. (2010); Mayer et al. (2009); Keys et al. (2012) Demyanyk and Van Hemert (2011); Purnanandam (2011); Dell'Ariccia et al. (2012); Piskorski et al. (2015); Griffin and Maturana (2016b); Griffin and Maturana (2016a)). The findings in this literature suggest that the subprime mortgage market was plagued with incentive problems, fraud, and poor underwriting, and represented a "classic lending boom-bust scenario" (Demyanyk and Van Hemert (2011)). As mentioned above, we focus more broadly on the PLS market; subprime mortgages make up at most 40% of originations in this market. To the best of our knowledge, this study is the first to isolate cross-sectional variation in geographic exposure to the acceleration of the PLS market in 2003, and it is the first to test how the rise of the PLS market affected house prices, construction activity, and housing market optimism.

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

This study is also related to the body of research exploring the role of speculation and

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

There is a body of research exploring the rise of household debt across the income and credit score distribution during the housing boom (Mian and Sufi (2009); Adelino et al. (2016); Mian and Sufi (2017a); Mian and Sufi (2017b); Adelino et al. (2017); Foote et al. (2016); Albanesi et al. (2017)). The focus of this research is on the rise in the level of household debt, which is not an outcome explored in this study. There is no claim in this study that the PLS market directly explains the aggregate rise in household debt, which was driven by existing homeowners extracting equity during the housing boom (e.g., Mian and Sufi (2011); Mian and Sufi (2015); Mian and Sufi (2017b)).

The analysis here does not focus on the average income or average credit score of zip codes or MSAs; instead, we define as our instrument exposure to the PLS market, which is positively correlated with income across MSAs, but negatively correlated within MSAs. Our results are consistent with some of the correlations in the existing literature. For example, we show that 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 our 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.[15]

[15]To the best of our knowledge, the only other study that constructs an instrument for credit supply for the 2002 to 2006 housing boom is Di Maggio and Kermani (2017). We have explored their instrument, and

8.2 Future work

We exploit cross-sectional variation in geographical areas across the United States in exposure to lenders that expand mortgage originations due to the acceleration of the PLS market. High PLS share areas witness strong relative growth in mortgage originations, volume, house prices, and construction. High PLS share MSAs are more likely to become "boom" MSAs: elastic housing supply areas that experience a simultaneous rise in both construction and house price growth. Perhaps most interestingly, individuals living in these MSAs on average become more pessimistic on the housing market during the boom. This suggests that the PLS market may have increased by house prices through a small cluster of optimistic individuals.

We are exploring this idea in ongoing work on this project. More specifically, we are collecting data on the prominence of investor purchases based on (a) the number of individuals taking on more than one first mortgage from TransUnion and (b) the fraction of purchases by individuals not living in the same zip code as the house from DataQuick. Our plan is to see whether investor purchases rise disproportionately in high NCL share zip codes, and whether these purchases are more likely to be financed with a mortgage. If our prediction is correct that investors are driving many of the origination and volume patterns documented above, then it would support the argument that the PLS market spurred a volume and price boom by shifting purchasing power to more optimistic agents living outside the zip codes in question.

it appears to be orthogonal to the PLS share we use here.

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

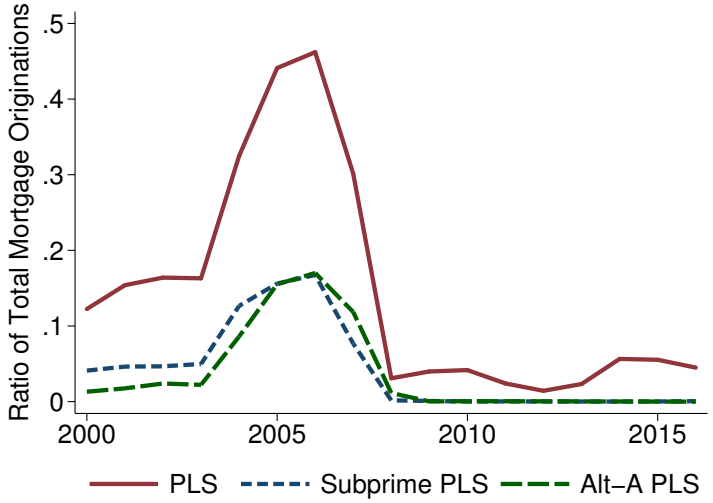
Piazzesi, Monika and Martin Schneider, "Momentum traders in the housing market: survey evidence and a search model," *American Economic Review*, 2009, *99* (2), 406–11.

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

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

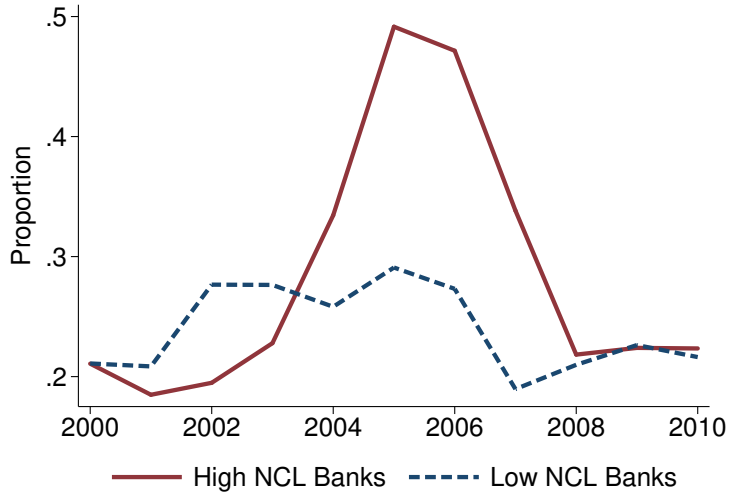
Saiz, Albert, "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

Figure 1: Acceleration of Private Label Securitization of Mortgages



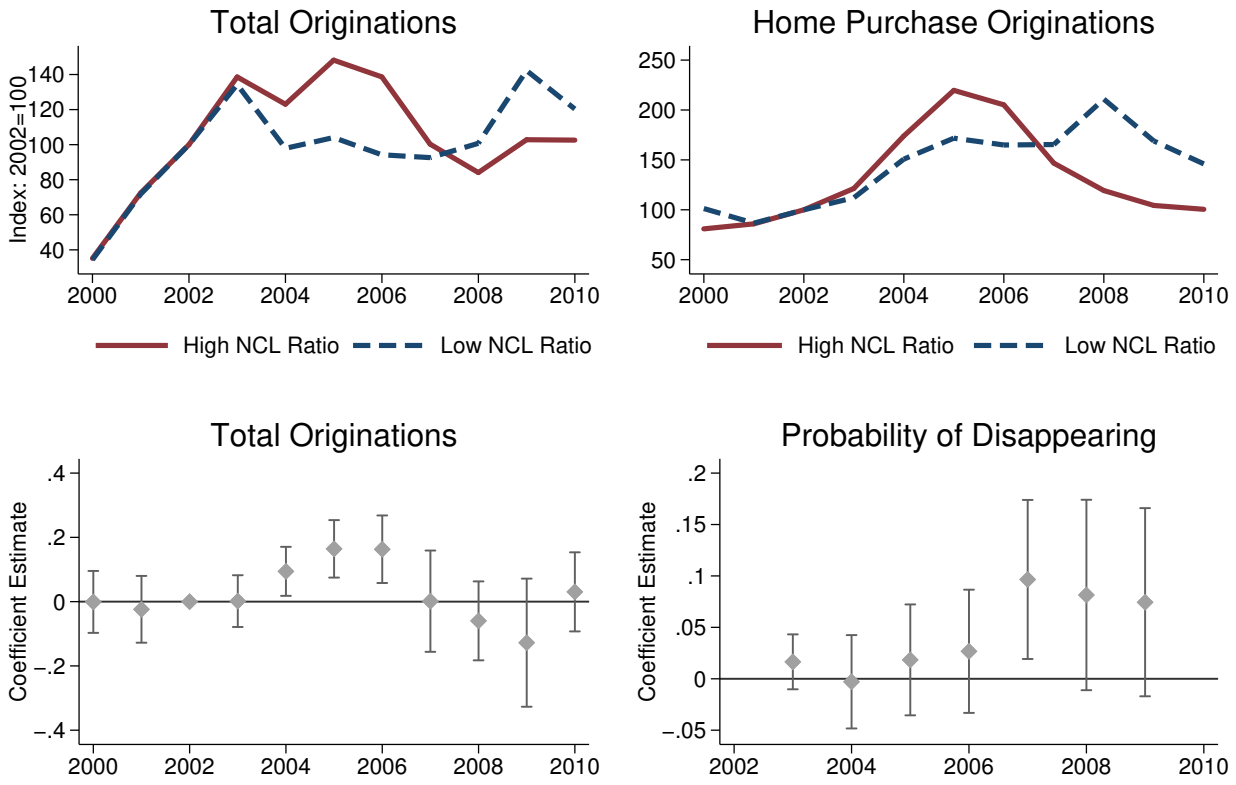
This figure plots the share of total mortgage originations that were private label securities (PLS), subprime PLS, and Alt-A PLS.

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



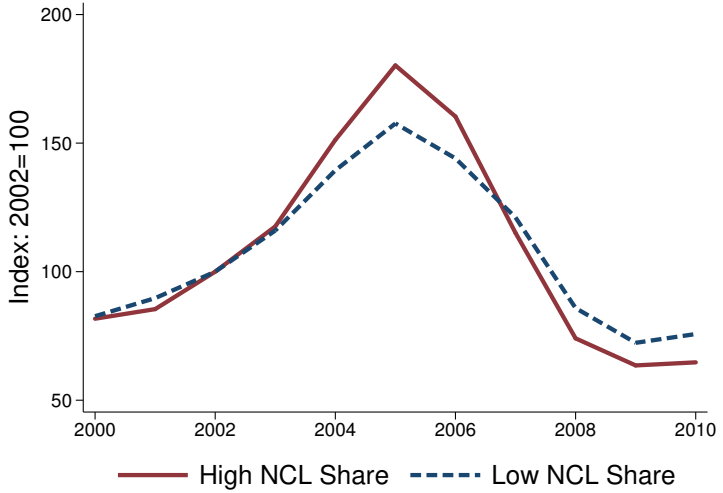
This figure plots the shares of total mortgage originations that were sold to private institutions by banks above and below the median non-core liabilities ratio (NCL) as of 2002. NCL is defined here as one minus the proportion of liabilities that are federally insured deposits for institutions that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

Figure 3: Relative Expansion of Originations by High NCL Ratio Lenders



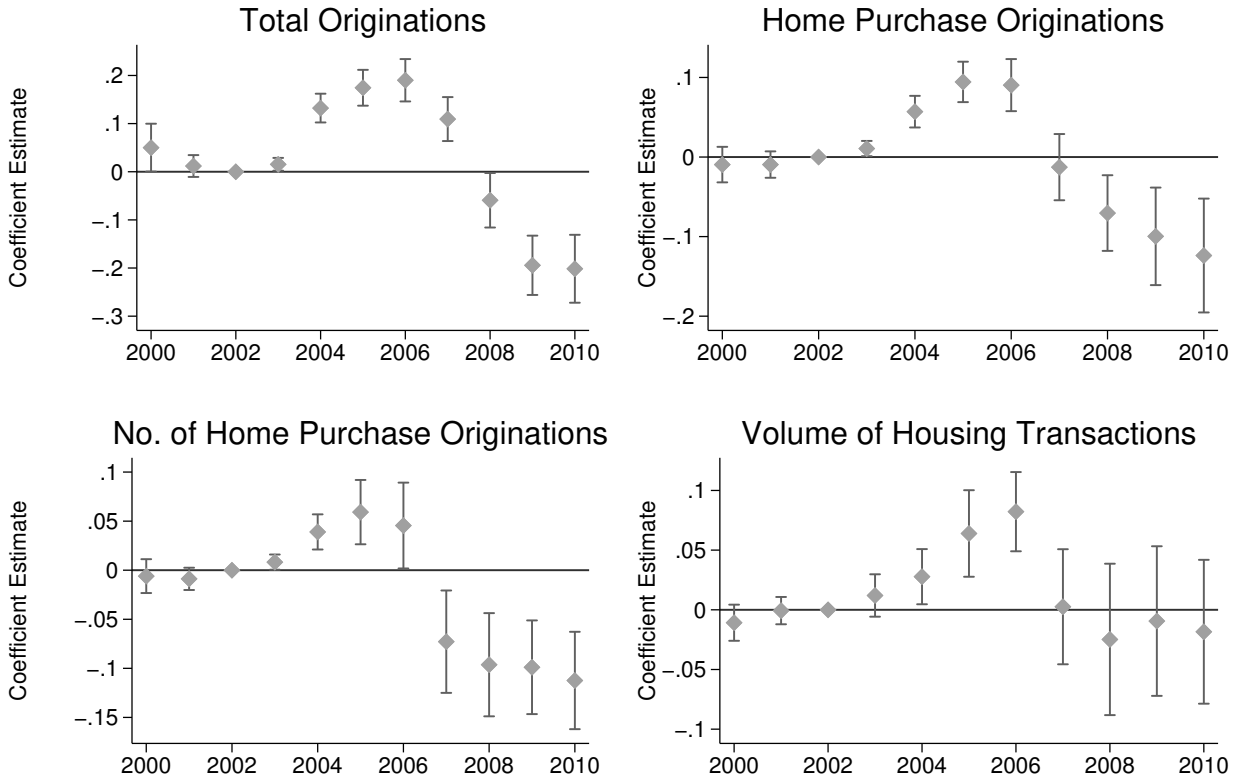
The top left panel plots the growth rates of mortgage originations of lenders above and below the median non-core liabilities ratio (NCL) as of 2002. The top right panel plots the growth rates of mortgage originations for home purchase of lenders above and below the median NCL as of 2002. The bottom left panel plots the coefficients $\{\beta_k\}$ of the specification $y_{b,t} = \alpha_b + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{b,2002} + \varepsilon_{b,t}$ for lender b at time t. $y_{b,t}$ is equal to total mortgage originations 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 here as one minus the proportion of liabilities that are federally insured deposits for institutions that are in the FFIEC Call Reports and one for institutions regulated by the Department of Housing and Urban Development (HUD). The regressions are weighted by the share of loans originated in 2002 by lender b. 95% confidence intervals from robust standard errors are also plotted. Lender fixed effects included in the panel regression.

Figure 4: Home Purchase Mortgage Originations in Zip Codes, by NCL Share of Lenders



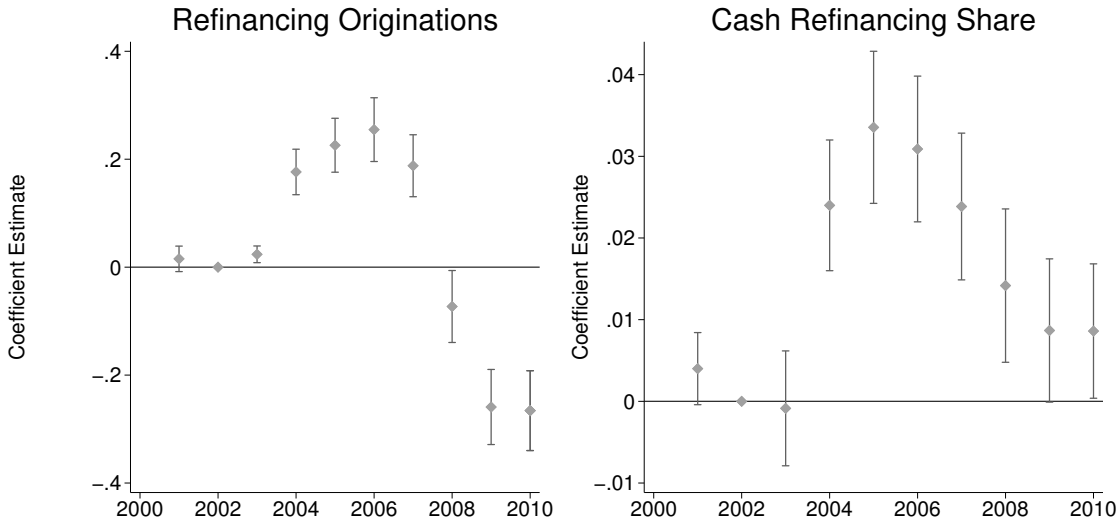
The figure plots the growth rates of mortgage originations for home purchases for zip codes above and below the median non-core liabilities share (NCL) as of 2002. NCL at the zipcode-level is defined as the weighted average of the NCL of lenders where the weights are the share of mortgages originated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

Figure 5: Zip-code Level Originations by NCL Share: Panel Regressions



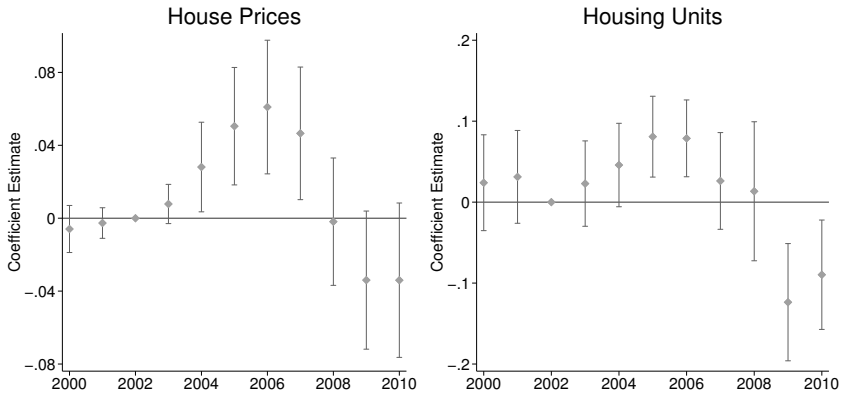
The panels plot the coefficients $\{\beta_k\}$ of the specification $y_{z,t} = \alpha_z + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{z,2002} + \varepsilon_{z,t}$ for zip code z in year t. $y_{z,t}$ in each panel is total mortgage originations (top left), total mortgage originations for home purchase (top right), the number of mortgage originations for home purchases (bottom left), and the volume of housing transactions (bottom right). 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 zipcode z. The regressions are weighted by the share of total occupied housing units in zip code z in 2000. 95% confidence intervals from robust standard errors are also plotted. Zip code level fixed effects included.

Figure 6: Zip-code Level Refinancing by NCL Share: Panel Regressions



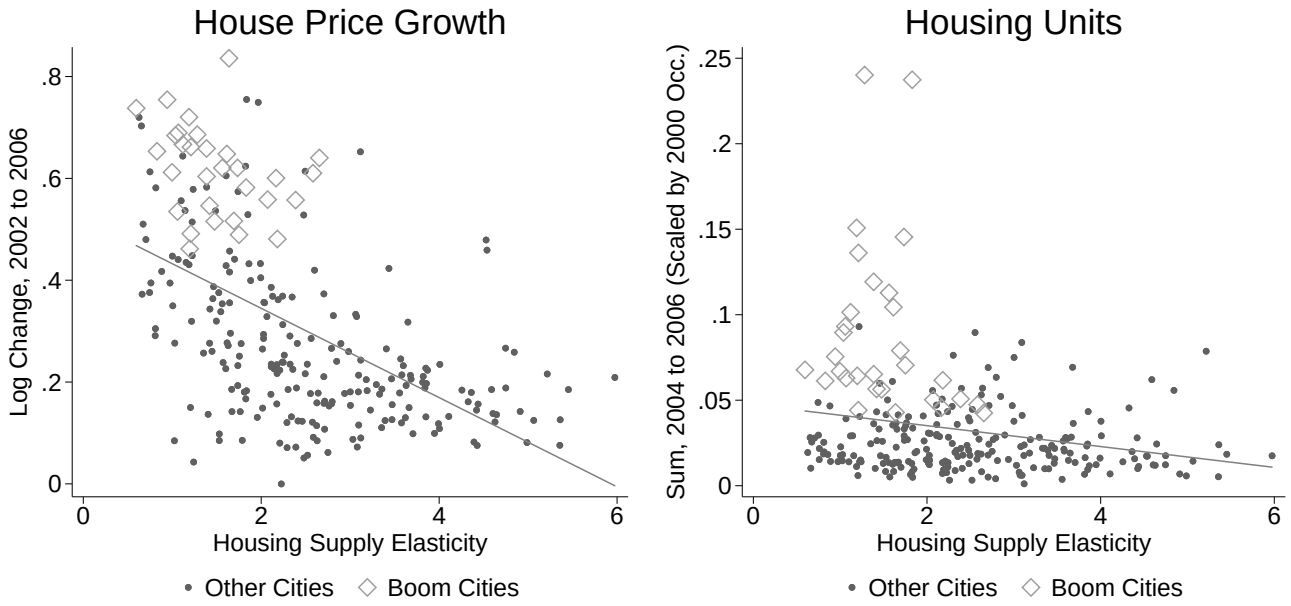
The panels plot the coefficients $\{\beta_k\}$ of the specification $y_{z,t} = \alpha_z + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{z,2002} + \varepsilon_{z,t}$ for zip code z in year t. $y_{z,t}$ in each panel is mortgage originations for refinancing (left) and share of existing mortgages that are refinanced with cash taken out (right). 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 zipcode 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 7: House Prices and Construction by NCL share: Panel Regression



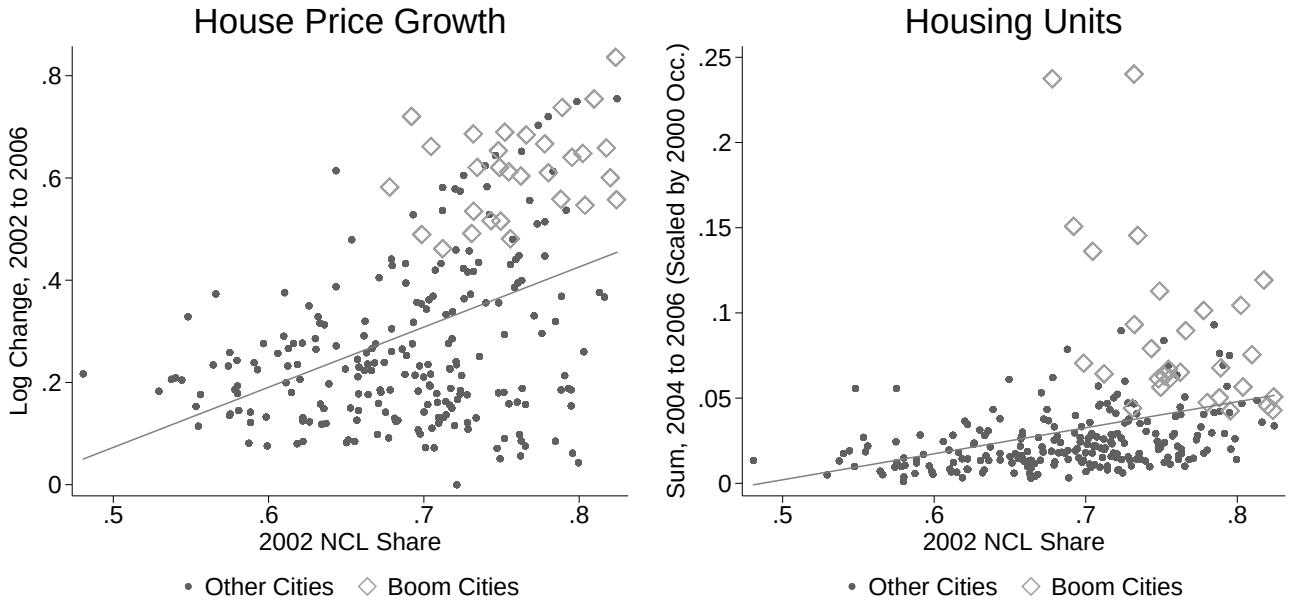
The left panel plots the coefficients $\{\beta_k\}$ of the specification $HP_{z,m,t} = \alpha_{z,m} + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{z,m,2002} + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k SAIZ_{m,2002} + \varepsilon_{z,m,t}$ for zipcode z, MSA m, and year t. The right panel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

Figure 8: The Housing Boom and Housing Supply Elasticity across MSAs



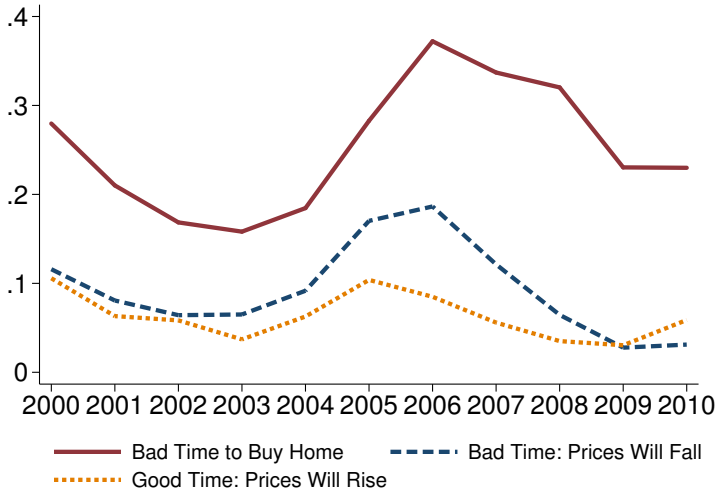
Plotted is the MSA-level correlation of the Saiz (2010) elasticity measure with house price growth from 2002 to 2006 (left panel) and the sum of housing units constructed from 2004 to 2006 scaled by the number of units in 2000 at the MSA level. Boom cities are defined to be those cities in the largest quartiles of both the house price growth and construction growth distribution.

Figure 9: The Housing Boom and the NCL Share across MSAs



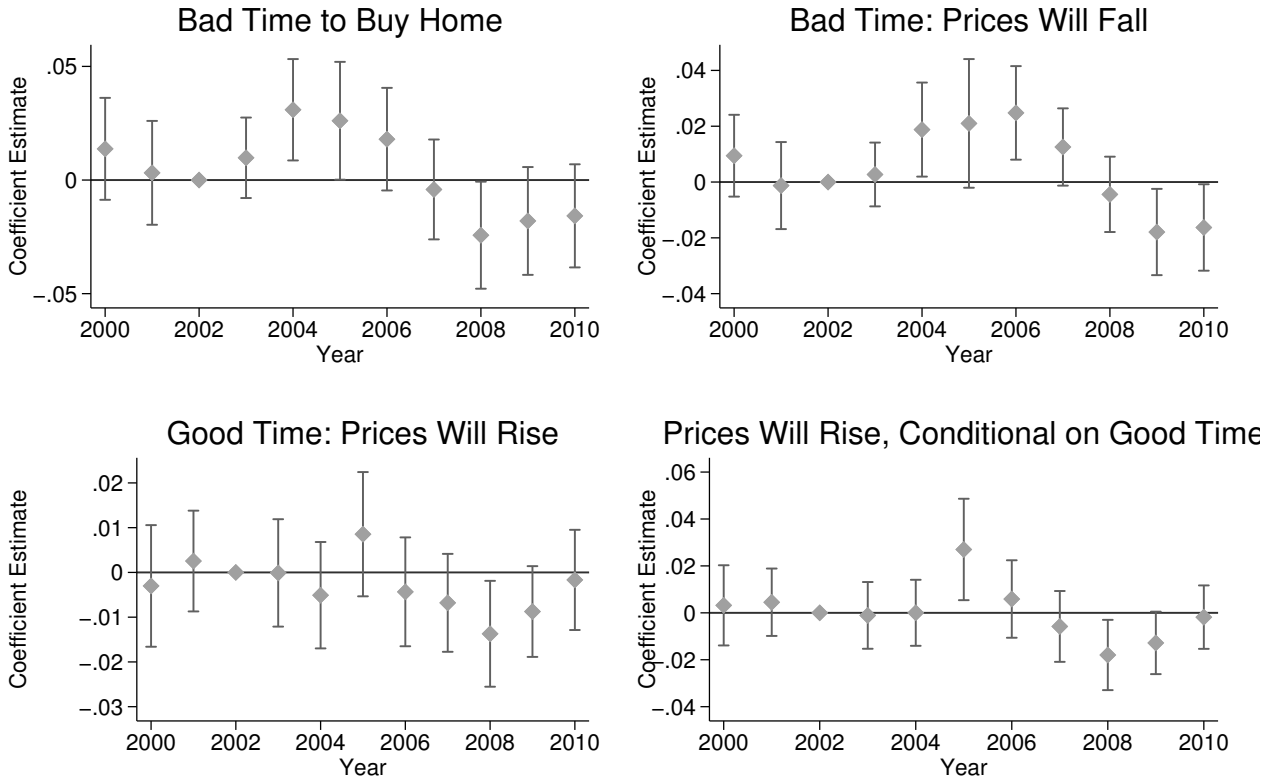
Plotted is the MSA-level correlation of the 2002 NCL share of the MSA with house price growth from 2002 to 2006 (left panel) and the sum of housing units constructed from 2004 to 2006 scaled by the number of units in 2000 at the MSA level. Boom cities are defined to be those cities in the largest quartiles of both the house price growth and construction growth distribution. The NCL share 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 zipcode z (MSA m).

Figure 10: Measures of Optimism on Housing Market from the Michigan Survey



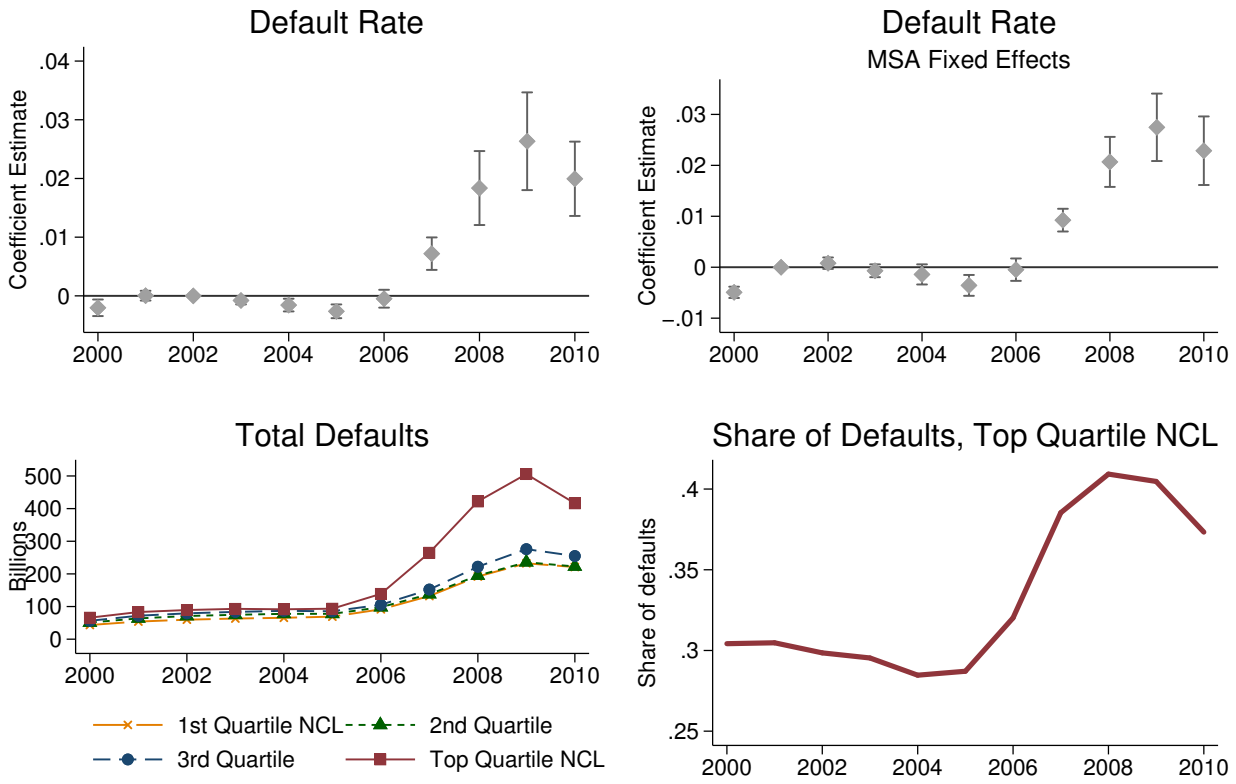
Plots the share of individuals that respond to the question "Good or bad time to buy home" with the answers "bad time to buy a home", "bad time because prices will fall", and "good time because prices will rise". See the text for more details.

Figure 11: Measures of Housing Market Optimism and the 2002 NCL Share: MSA-level



The panels plot the coefficients $\{\beta_k\}$ of the specification $y_{m,t} = \alpha_m + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{m,2002} + \varepsilon_{m,t}$ for MSA m in year t. $y_{m,t}$ in each panel is "bad time to buy a home" (top left), "bad time to buy a home because prices will fall" (top right), "good time to buy a home because prices will rise" (bottom left), and the share that answered "good time to buy a home because prices will rise" conditional on first answering it is a "good time to buy a home" (bottom right). 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 Michigan survey respondents in zipcode m in 2002. 95% confidence intervals from robust standard errors are also plotted. MSA level fixed effects included.

Figure 12: Household Debt Defaults by NCL Share



The top panels plot the coefficients $\{\beta_k\}$ of the specification $DEF_{z,t} = \alpha_z + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{z,2002} + \varepsilon_{z,t}$ for zipcode z in year t. $DEF_{z,t}$ is the default rate in zipcode z at year t. The top left panel is without MSA fixed effects and the top right panel is with MSA fixed effects. 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 zipcode 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. The bottom left panel plots total defaults in zip codes by $NCL_{z,2002}$ population-weighted quartiles. The bottom right panel plots the share of total defaults in zip codes in the highest NCL exposure quartile.

Table 1: Summary Statistics

Panel A: Lender-Level Data

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

Panel B: Zip-Code-Level Data

	count	mean	p50	sd
2002 NCL Share	10865	0.78	0.78	0.04
Δ_{BOOM} ln (Loans Originated)	10862	0.44	0.37	0.43
Δ_{BOOM} ln (Loans Originated for Home Purchase)	10860	0.57	0.54	0.35
Δ_{BOOM} ln (Loans Originated for Refinancing)	10848	0.32	0.22	0.52
Δ_{BOOM} ln (Number of Loans Originated)	10862	0.15	0.11	0.32
Δ_{BOOM} ln (Number of Loans Originated for Home Purchase)	10860	0.29	0.26	0.28
Δ_{BOOM} ln (Number of Loans Originated for Refinancing)	10848	0.02	-0.04	0.39
$\Delta_{02,05}$ Cash Refinancing Share	10846	-0.01	-0.01	0.08
Δ_{BOOM} ln (Volume of Housing Transactions)	4219	0.25	0.19	0.35
$\Delta_{02,06}$ ln (House Prices)	5805	0.37	0.36	0.22

Panel C: MSA-Level Data

	count	mean	p50	sd
2002 NCL Share	259	0.71	0.72	0.06
Housing Supply Elasticity	259	1.96	1.65	1.18
$\sum_{04-06}$ Housing Units per Occupants in 2000	259	0.03	0.02	0.03

Table 2: High NCL Ratio Predicts Growth in Mortgage Originations

	Change in Fraction of PLS, 2002-05	Originations Growth, 2002 to 2005			Originations Growth, Pre-Boom	
	(1)	(2)	(3)	(4)	(5) 1998-2000	(6) 2000-2002
2002 NCL Ratio	0.151**	0.183***		0.203*	-0.00641	-0.0828
	(0.049)	(0.047)		(0.082)	(0.046)	(0.105)
Non-Bank 2002			0.284*	-0.0665		
			(0.114)	(0.189)		
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 is in column 1. Regression results for the specification $\Delta HMDA_{b,p} = \alpha + \beta_0 NCL_{b,2002} + \beta_1 NB_{b,2002} + \varepsilon_{b,p}$ for lender b in period $p \in \{BOOM, PRE\}$ are in columns 2-5. $\Delta PLS_{b,2002,2005}$ is the change in the share of loan originations that were privately securitized by lender b from 2002 to 2005, $\Delta HMDA_{b,BOOM}$ is the growth in loan orginations by lender b from 2000-2002 to 2004-2006, and $\Delta HMDA_{b,PRE}$ is the growth in loan orginations by lender b from 1998-1999 to 2001-2002 . NCL is defined here as one minus the proportion of liabilities that are federally insured deposits for institutions that are in the FFIEC Call Reports and one for institutions regulated by the Department of Housing and Urban Development (HUD). A non-bank mortgage lender is an institution regulated by the HUD. Column 1 is restricted to commercial banks and thrifts. Standard errors are robust.

Table 3: NCL Share Correlations with Observatble Variables

Covariates	Across MSA	Within MSA
2000 Deposits/Loan Originations	-.091***	
	(.017)	
Saiz Elasticity	-.374***	
	(.073)	
2000 Fraction Age 65+	-.007***	-.010***
	(.002)	(.000)
2000 Fraction Hispanic or Black	.067***	.118***
	(.010)	(.003)
2000 Fraction Renters	.019*	.039***
	(.007)	(.002)
2000 Fraction Subprime	.018***	.067***
	(.003)	(.001)
2000 Log Median Home Value	.116***	-.068***
	(.024)	(.007)
2000 Log Median Household Income	.021	-.100***
	(.014)	(.005)

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

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

	Bank-MSA Originations		Bank-Zip-Code Originations	
	(1)	(2)	(3)	(4)
2002 NCL Ratio	0.169***	0.141***	0.177***	0.123***
	(0.007)	(0.006)	(0.003)	(0.002)
Geography FE	N	Y	N	Y
N	65107	65107	714358	714358
R-sq	0.041	0.161	0.032	0.363

Standard errors in parentheses

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

Regression results for the specification $\Delta y_{b,m,BOOM} = \alpha_m + \beta NCL_{b,2002} + \varepsilon_{b,m,BOOM}$ for lender b in MSA m are in columns 1-2. Regression results for the specification $\Delta y_{b,z,BOOM} = \alpha_z + \beta NCL_{b,2002} + \varepsilon_{b,z,BOOM}$ for lender b in zipcode z are in columns 3-4. $\Delta y_{b,m,BOOM}$ ($\Delta y_{b,z,BOOM}$) is defined here as total mortgage originations growth from 2000-2002 to 2004-2006 for lender b in MSA m (zipcode 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 (zipcode z). Standard errors are clustered at the MSA (zipcode) level.

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

	Originations		Purchase Originations		No. of Purchase Originations		Volume	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2002 NCL Share	0.153***	0.246***	0.0889***	0.142***	0.0547***	0.154***	0.0637***	0.0848*
	(0.021)	(0.023)	(0.014)	(0.018)	(0.013)	(0.018)	(0.015)	(0.039)
MSA FE	N	Y	N	Y	N	Y	N	Y
N	10862	10862	10860	10860	10860	10860	4219	4219
R-sq	0.129	0.621	0.063	0.438	0.038	0.353	0.031	0.261

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 zipcode z and MSA m. The outcomes are mortgage originations growth, mortgage orgiinations for home purchase growth, number of originations for home purchase growth, and volume of housing transactions growth from 2000-2002 to 2004-2006. The NCL at the zipcode-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 zipcode z. Regressions are weighted by the share of occupied houselds in zipcode z and year 2000. Standard errors are clustered at the MSA-Level.

Table 6: NCL Share and Refinancings: Zip Code Level

	Refinancing Originations		Cash Refinancing Share	
	(1)	(2)	(3)	(4)
2002 NCL Share	0.204***	0.337***	0.0335***	0.0352***
	(0.027)	(0.030)	(0.005)	(0.004)
MSA FE	N	Y	N	Y
N	10848	10848	10846	10846
R-sq	0.153	0.683	0.187	0.691

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 zipcode z and MSA m. The outcomes are mortgage originations for refinancing growth from 2000-2002 to 2004-2006 and the change in cash refinancing shares from 2002 to 2005. The NCL at the zipcode-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 zipcode z. Regressions are weighted by the share of occupied houselds in zipcode z and year 2000. Standard errors are clustered at the MSA-Level.

Table 7: NCL Share and House Price Growth

	House Price Growth, 2002 to 2006					
	(1)	(2)	(3)	(4)	(5)	(6)
Housing Supply Elasticity	-0.121***		-0.110***	0.761**		0.000
	(0.019)		(0.014)	(0.278)		(.)
2002 NCL Share		0.079***	0.061***	0.121***	0.027**	0.068***
		(0.024)	(0.018)	(0.029)	(0.008)	(0.015)
2002 NCL Share X Housing Supply Elasticity				-0.042**		-0.027***
				(0.014)		(0.006)
MSA FE	N	N	N	N	Y	Y
N	4834	4834	4834	4834	4834	4834
R-sq	0.298	0.132	0.374	0.405	0.930	0.933

Standard errors in parentheses

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

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

Table 8: NCL Share and Construction

	Housing Units 2004-2006			
	(1)	(2)	(3)	(4)
Housing Supply Elasticity	-0.00136		0.00269	0.00302
	(0.002)		(0.002)	(0.002)
2002 NCL Share		0.0117***	0.0128***	0.0111*
		(0.002)	(0.002)	(0.005)
2002 NCL Share X Housing Supply Elasticity				0.000810
				(0.002)
N	259	259	259	259
R-sq	0.003	0.181	0.193	0.194

Standard errors in parentheses

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

Regression results for the specification $UNITS_{m,2004,2006} = \alpha + \beta SAIZ_m + \varepsilon_m$ from MSA m is in column 1. Regression results for the specification $UNITS_{m,2004,2006} = \alpha + \beta NCL_{m,2002} + \varepsilon_m$ from MSA m is in column 2. Regression results for the specification $UNITS_{m,2004,2006} = \alpha + \beta_0 SAIZ_m + \beta_1 NCL_{m,2002} + \varepsilon_m$ from MSA m is in column 3. Regression results for the specification $UNITS_{m,2004,2006} = \alpha + \beta_0 SAIZ_m + \beta_1 NCL_{m,2002} + \beta_2 NCL_{m,2002} SAIZ_m + \varepsilon_m$ from MSA m is in column 4. $UNITS_{m,2004,2006}$ is the sum of housing units from 2004 to 2006 in MSA m and $SAIZ_m$ 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 9: NCL Share and Boom MSAs

	In Highest HP and Units Quartiles during Boom			House Price Growth, 2002-2010		
	(1)	(2)	(3)	(4)	(5)	(6)
Housing Supply Elasticity	-0.0515**		-0.0209	-0.0110		-0.0191
	(0.018)		(0.015)	(0.010)		(0.011)
2002 NCL Share		0.106***	0.0971***		-0.0192	-0.0269
		(0.029)	(0.028)		(0.013)	(0.014)
N	259	259	259	253	253	253
R-sq	0.039	0.121	0.126	0.007	0.015	0.032

Standard errors in parentheses

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

Regression results for the specification $y_m = \alpha + \beta SAIZ_m + \varepsilon_m$ from MSA m are in columns 1 and 4. Regression results for the specification $y_m = \alpha + \beta NCL_{m,2002} + \varepsilon_m$ from MSA m are in columns 2 and 5. Regression results for the specification $y_m = \alpha + \beta_0 SAIZ_m + \beta_1 NCL_{m,2002} + \beta_2 NCL_{m,2002} SAIZ_m + \varepsilon_m$ from MSA m are in columns 3 and 6. For columns 1-3, y_m is an indicator equal to 1 when an MSA's 2002-2006 housing price growth and 2004-2006 sum of housing units are each in the highest quartile of their distirbutions and 0 otherwise; y_m is housing price growth from 2002 to 2010 in columns 4-6. $SAIZ_m$ 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 10: NCL Share and Housing Market Optimism: MSA Level

	(1) $\Delta_{00,02}$ Good: Prices	(2) $\Delta_{00,02}$ ICE	(3) Δ_{BOOM} Bad	(4) Δ_{BOOM} Bad: Prices	(5) Δ_{BOOM} Good: Prices	(6) Δ_{BOOM} ICE
2002 NCL Share	0.00126	0.00878	0.0239*	0.0222*	0.000565	0.0100
	(0.007)	(0.012)	(0.010)	(0.009)	(0.005)	(0.008)
N	237	237	244	244	244	244
R-sq	0.000	0.002	0.043	0.053	0.000	0.006

Standard errors in parentheses

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

Regression results for the specification $y_m = \alpha + \beta NCL_{m,2002} + \varepsilon_m$ from MSA m are in columns 1-6. The outcomes y_m are the 2000-2002 change in the share of respondents answering it is a good time to purchase a home because prices will rise; the 2000-2002 change in the Index of Consumer Expectations (ICE); the 2004-2006 average share of respondents saying it is a bad time to purchase a home minus the 2000-2003 average share; the 2004-2006 average share of respondents saying it is a bad time to purchase a home because prices will fall minus the 2000-2003 average share; the 2004-2006 average share of respondents saying it is a good time to purchase a home because prices will rise minus the 2000-2003 average share; and the 2004-2006 average ICE minus that of 2000-2003, respectively. The regressions are weighted by the share of survey participants in MSA m in 2002. Standard errors are robust.

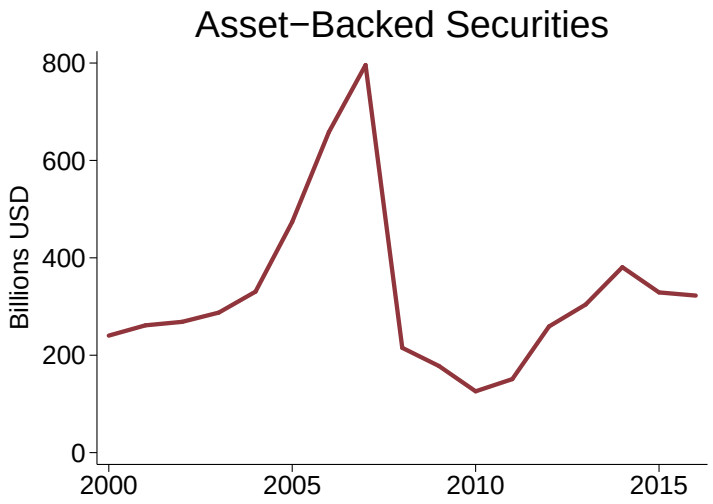
Credit Supply, Beliefs, and Speculation: Private Label Securitization and the Housing Cycle of 2000 to 2010 Appendix

Atif Mian
Princeton University and NBER

Amir Sufi
University of Chicago Booth School of Business and NBER

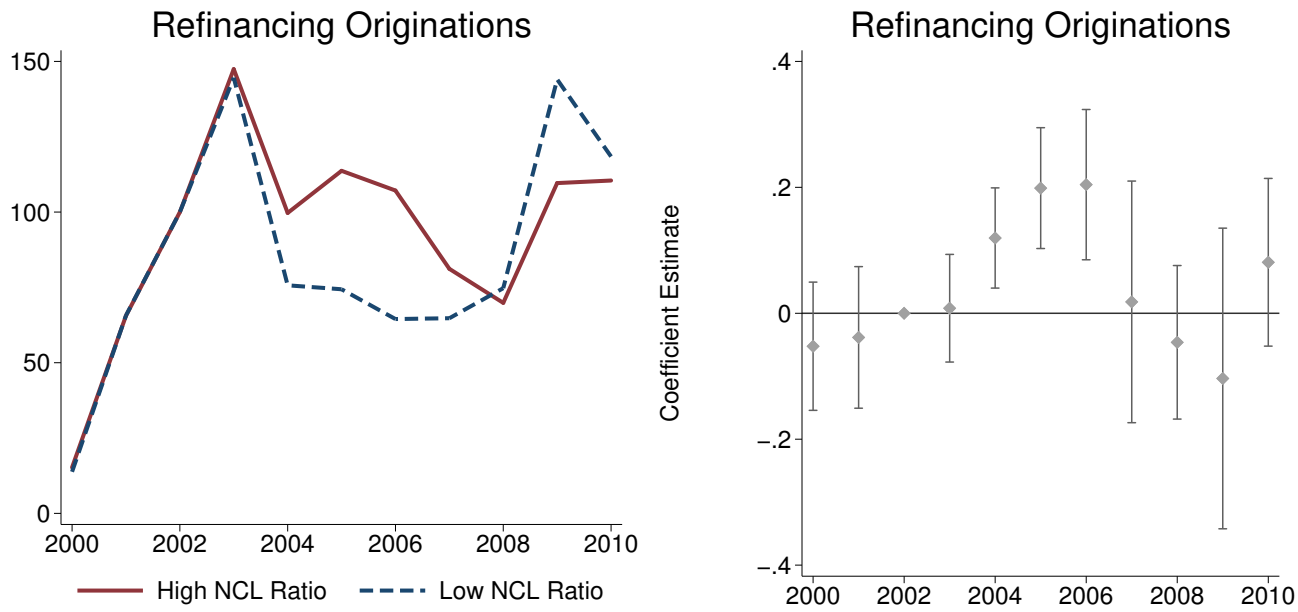
February 2018

Figure 1: Rise in Non-Mortgage ABS



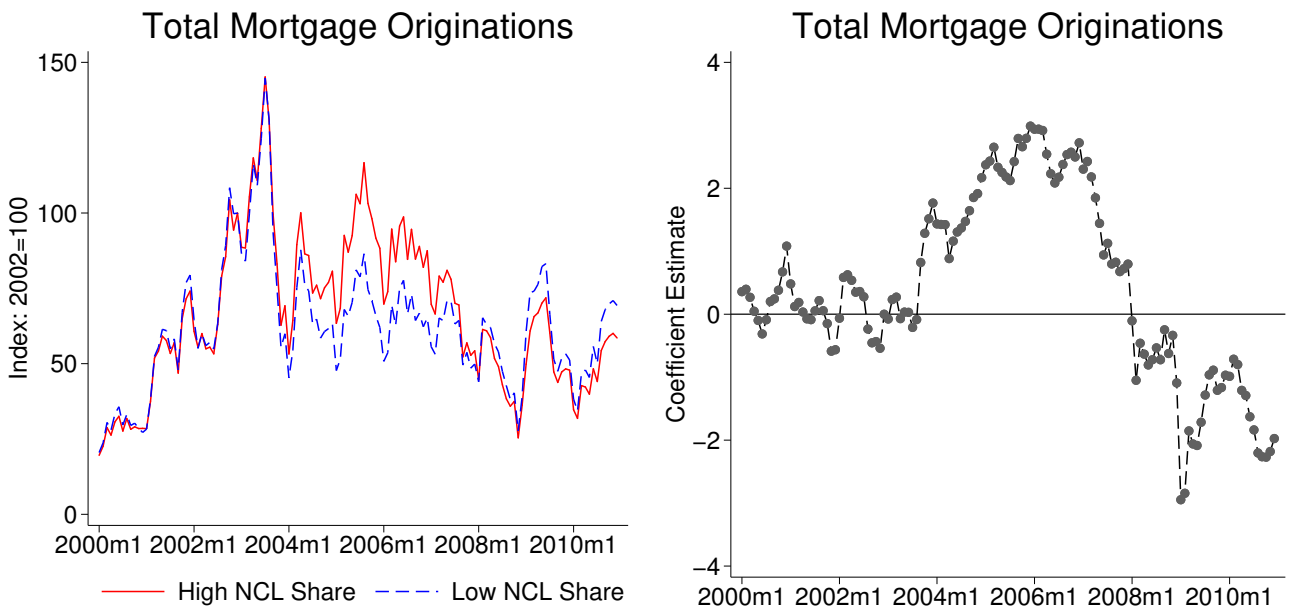
Plotted is the amount of asset-backed securities in billions of US dollars.

Figure 2: Lender Level Results for Refinancing Originations



The left panel plots the growth rates of mortgage originations for refinancing of lenders above and below the median non-core liabilities share (NCL). The right panel plots the coefficients $\{\beta_k\}$ of the specification $y_{b,t} = \alpha_b + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{b,2002} + \varepsilon_{b,t}$ for lender b at time t. $y_{b,t}$ is equal to total mortgage originations for refinancing by a lender b in year t. NCL is defined here as one minus the proportion of liabilities that are federally insured deposits for institutions that are in the FFIEC Call Reports and one for institutions regulated by the Department of Housing and Urban Development (HUD). The regressions are weighted by the share of refinancing loans originated in 2002 by lender b. 95% confidence intervals from robust standard errors are also plotted. Lender fixed effects included in panel regression.

Figure 3: MSA Monthly Level Results for Mortgage Originations



The left panel plots the growth rates of mortgage originations for MSAs above and below the median non-core liabilities share (NCL). The right panel plots the coefficients $\{\beta_k\}$ of the specification $y_{m,t} = \alpha_m + \gamma_t + \sum_{k \neq 2002} \mathbb{1}_{t=k} \beta_k NCL_{m,2002} + \varepsilon_{m,t}$ for MSA m at time t. $y_{m,t}$ is equal to total mortgage originations by MSA m in year t. 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. 95% confidence intervals from robust standard errors are also plotted. Lender fixed effects included in panel regression.

Table 1: NCL Share and Mortgage Origination Growth: Zip Code Level

	Originations		Purchase Originations		No. of Purchase Originations		Volume	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1998 NCL Share	0.157***	0.262***	0.0976***	0.166***	0.0778***	0.164***	0.0768***	0.0952*
	(0.021)	(0.017)	(0.015)	(0.015)	(0.011)	(0.013)	(0.013)	(0.038)
MSA FE	N	Y	N	Y	N	Y	N	Y
N	10856	10856	10855	10855	10855	10855	4213	4213
R-sq	0.136	0.668	0.077	0.475	0.077	0.394	0.051	0.278

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 zipcode z and MSA m. The outcomes are mortgage originations growth, mortgage orgiinations for home purchase growth, number of originations for home purchase growth, and volume of housing transactions growth from 2000-2002 to 2004-2006. The NCL at the zipcode-level is defined as the weighted average of NCL at the bank-level where the weights are the share of loans originated in 2002 by a bank b in zipcode z. Regressions are weighted by the share of occupied houselds in zipcode z and year 2000. Standard errors are clustered at the MSA-Level.

Table 2: NCL Share and Refinancings: Zip Code Level

	Refinancing Originations		Cash Refinancing Share	
	(1)	(2)	(3)	(4)
1998 NCL Share	0.200***	0.343***	0.0330***	0.0338***
	(0.028)	(0.021)	(0.005)	(0.005)
MSA FE	N	Y	N	Y
N	10843	10843	10840	10840
R-sq	0.147	0.726	0.181	0.702

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 zipcode z and MSA m. The outcomes are mortgage originations for refinancing growth from 2000-2002 to 2004-2006 and the change in cash refinancing shares from 2002 to 2005. The NCL at the zipcode-level is defined as the weighted average of NCL at the bank-level where the weights are the share of loans originated in 2002 by a bank b in zipcode z. Regressions are weighted by the share of occupied houselds in zipcode z and year 2000. Standard errors are clustered at the MSA-Level.

Table 3: NCL Share and House Price Growth

	House Price Growth, 2002 to 2006					
	(1)	(2)	(3)	(4)	(5)	(6)
Housing Supply Elasticity	-0.121***		-0.118***	0.563**		0.000
	(0.019)		(0.016)	(0.201)		(.)
1998 NCL Share		0.061**	0.054**	0.115***	0.027***	0.060***
		(0.022)	(0.017)	(0.030)	(0.007)	(0.013)
1998 NCL Share X Housing Supply Elasticity				-0.042**		-0.023***
				(0.013)		(0.006)
MSA FE	N	N	N	N	Y	Y
N	4834	4828	4828	4828	4828	4828
R-sq	0.298	0.073	0.355	0.383	0.931	0.934

Standard errors in parentheses

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

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

Table 4: NCL Share and Construction

	Housing Units 2004-2006			
	(1)	(2)	(3)	(4)
Housing Supply Elasticity	-0.00136		0.00196	0.00216
	(0.002)		(0.002)	(0.002)
1998 NCL Share		0.0114***	0.0121***	0.00989*
		(0.002)	(0.002)	(0.004)
1998 NCL Share X Housing Supply Elasticity				0.00110
				(0.002)
N	259	259	259	259
R-sq	0.003	0.172	0.179	0.181

Standard errors in parentheses

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

Regression results for the specification $UNITS_{m,2004,2006} = \alpha + \beta SAIZ_m + \varepsilon_m$ from MSA m is in column 1. Regression results for the specification $UNITS_{m,2004,2006} = \alpha + \beta NCL_{m,2002} + \varepsilon_m$ from MSA m is in column 2. Regression results for the specification $UNITS_{m,2004,2006} = \alpha + \beta_0 SAIZ_m + \beta_1 NCL_{m,2002} + \varepsilon_m$ from MSA m is in column 3. Regression results for the specification $UNITS_{m,2004,2006} = \alpha + \beta_0 SAIZ_m + \beta_1 NCL_{m,2002} + \beta_2 NCL_{m,2002} SAIZ_m + \varepsilon_m$ from MSA m is in column 4. $UNITS_{m,2004,2006}$ is the sum of housing units from 2004 to 2006 in MSA m and $SAIZ_m$ 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 5: NCL Share and Boom MSAs

	In Highest HP and Units Quartiles during Boom			House Price Growth, 2002-2010		
	(1)	(2)	(3)	(4)	(5)	(6)
Housing Supply Elasticity	-0.0515**		-0.0259	-0.0110		-0.0158
	(0.018)		(0.015)	(0.010)		(0.011)
1998 NCL Share		0.103***	0.0934***		-0.0132	-0.0187
		(0.029)	(0.027)		(0.013)	(0.014)
N	259	259	259	253	253	253
R-sq	0.039	0.114	0.123	0.007	0.007	0.020

Standard errors in parentheses

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

Regression results for the specification $y_m = \alpha + \beta SAIZ_m + \varepsilon_m$ from MSA m are in columns 1 and 4. Regression results for the specification $y_m = \alpha + \beta NCL_{m,2002} + \varepsilon_m$ from MSA m are in columns 2 and 5. Regression results for the specification $y_m = \alpha + \beta_0 SAIZ_m + \beta_1 NCL_{m,2002} + \beta_2 NCL_{m,2002} SAIZ_m + \varepsilon_m$ from MSA m are in columns 3 and 6. For columns 1-3, y_m is an indicator equal to 1 when an MSA's 2002-2006 housing price growth and 2004-2006 sum of housing units are each in the highest quartile of their distirbutions and 0 otherwise; y_m is housing price growth from 2002 to 2010 in columns 4-6. $SAIZ_m$ 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.