Did Dubious Mortgage Origination Practices Distort House Prices?

John M. Griffin
McCombs School of Business, University of Texas at Austin

Gonzalo Maturana
Goizueta Business School, Emory University

ZIP codes with high concentrations of originators who misreported mortgage information experienced a 75% larger relative increase in house prices from 2003 to 2006 and a 90% larger relative decrease from 2007 to 2012 compared with other ZIP codes. Several causality tests show that high fractions of dubious originators in a ZIP code lead to large price distortions. Originators with high misreporting gave credit to borrowers with high ex ante risk, yet further understated the borrowers' true risk. Overall, excess credit facilitated through dubious origination practices explain much of the regional variation in house prices over a decade. (*JEL* G21, G23, R30, R31)

Received August 24, 2015; accepted January 22, 2016 by Editor Matthew Spiegel.

Are the costs of fraud large or small? Akerlof and Romer (1993) discuss how financial corruption can cause aggregate price distortions that are much larger than the amounts gained from the dubious activity.[1] They point to the U.S. savings and loans crisis as an example in which developers and bankers extracted rents from thrifts through nonrecourse construction loans. However, the combined activity had the unintentional effect of at least amplifying a commercial real estate building boom and an ultimate bust. In a similar vein, but

We are grateful to anonymous referees, Manuel Adelino, Andres Almazan, Aydogan Alti, Itzhak Ben-David, Carlos Carvalho, Jonathan Cohn, Cesare Fracassi, Jay Hartzell, Nicholas Hirschey, Wei Jiang, Sam Kruger, Elena Loutskina, Atif Mian, Jordan Nickerson, Amiyatosh Purnanandam, Konark Saxena, Mark Seasholes, Amit Seru, Amin Shams, Cindy Soo, Denis Sosyura, Matthew Spiegel (the editor), Mike Stegemoller, Parth Venkat, Vikrant Vig, and seminar participants at the 2015 AREUEA Conference, the 2014 Australasian Finance and Banking Conference, the 2015 SFS Cavalcade, Baylor University, Georgia State University, London Business School, London School of Economics, Southern Methodist University, University of Michigan, the University of New South Wales, the University of Texas at Austin, the University of Texas at Dallas, and the University of Virginia for helpful comments. The data used in this paper were purchased by John Griffin through Integra FEC. Griffin is an owner of Integra FEC and Integra REC, which engage in financial consulting on a variety of issues, including mortgage fraud. Integra REC is a party to MBS litigation against banks. The analyses and conclusions of this paper are unaffected. Supplementary results can be found in an Internet Appendix on the *Review of Financial Studies* Web site. Send correspondence to Gonzalo Maturana, Emory University, Department of Finance, Goizueta Business School, 1300 Clifton Road, Atlanta, GA 30320; telephone: 404-727-7497. E-mail: gonzalo.maturana@emory.edu.

[1] More generally, the potential distortive costs of bad financing have been recently summarized and emphasized by Zingales (2015).

doi:10.1093/rfs/hhw013 Advance Access publication March 15, 2016

through a different mechanism, we ask whether questionable origination practices led to any distortions in the recent 2003-2012 real estate boom and bust.

Misreporting a few features on a borrower's mortgage application may seem harmless in isolation; however, the process could extend credit to a borrower who may have little financial wherewithal or desire to repay. What if misreporting was not isolated but instead differed widely across loan originators with different geographic locales? What if the rise of securitization allowed these misreporting originators to issue many more loans? ZIP codes that contained a high presence of originators with questionable origination practices may have had relatively more undeserved loans than ZIP codes with better originators. This excess credit may have led to increased housing demand, causing a rise in prices and a subsequent crash in prices when the credit was removed.

We test the hypothesis that it was not securitization alone, but securitization coupled with bad originator practices that had the most distortive effects on house prices. Mian and Sufi (2009) show that subprime ZIP codes experienced a large increase in credit from 2002 to 2005 unrelated to income growth. This increase in credit can be traced to the rise of securitization (Nadauld and Sherlund 2013), providing support for a loan supply-based explanation for the 2002 to 2006 house price bubble. More recently, Adelino, Schoar, and Severino (2015, 2016) argue that Mian and Sufi's (2009) interpretation is incorrect. Our findings directly weigh in this debate. We find strong evidence that contradicts the demand view of Adelino, Schoar, and Severino and is generally consistent with the Mian and Sufi supply-based explanation. Rather than securitization alone, originator malfeasance in certain localities raised the credit supply; this drove up house prices relative to other areas and subsequently led to larger price crashes.

There is growing awareness that originators may have moved beyond lacking careful monitoring to actively pushing loans that did not meet underwriting standards. Ben-David (2011), Carrillo (2011), Garmaise (2015), and Jiang, Nelson, and Vytlacil (2014a) document appraisal, assets, and income misreporting at certain banks or geographic areas. Piskorski, Seru, and Witkin (2015) and Griffin and Maturana (2016) use different data and methods but find similar levels of second-lien misreporting spread across the nonagency securitization market. They find that second-lien misreporting varies widely across originators. This wide variation in misreporting practices across originators lays the empirical groundwork to ask if originators who engaged in large amounts of misreporting distorted house prices.

We classify originators in the highest tercile of second-lien misreporting in Griffin and Maturana (2016) as the "worse" (or "dubious") originators.[2] We measure the fraction of all transactions within a ZIP code by the worse, medium,

[2] Griffin and Maturana (2016) document that owner occupancy misreporting and appraisal inflation do not vary much across lenders but are primarily forms of misreporting attributed to borrowers and appraisers.

and better originators.[3] We first document that ZIP codes with high worse originator activity exhibited a larger rise in house prices from 2003 to 2006 and a larger decrease from 2007 to 2012. This relation is much stronger than the relation to prices for the fraction of nonagency loans securitized, and it holds within metropolitan statistical areas (MSAs) and after controlling for ZIP code income levels and income growth. The relation also holds within five quintiles based on 2001 income, indicating that the effect is not confined to subprime ZIP codes, yet the effects of fraud are strongest there. Overall, house prices in the 858 ZIP codes with the highest fraction of worse originators' market share increased 75% more (63% relative to 36%) over the 2003 to 2006 boom relative to the 4,318 ZIP codes with a lower presence of the worse originators. Conversely, from 2007 to 2012, these same 858 ZIP codes experienced a decrease in house prices nearly twice as large as the other ZIP codes (40% relative to 21%).

The strong relation between house prices and dubious origination practices in the ZIP codes need not be causal. To investigate causality, we take several approaches. First, we test whether originators with bad practices may simply be more aggressive at expanding into areas of rapidly increasing prices. Within an MSA, we find two ZIP codes with similar price run-up but different levels of dubious origination. Consistent with dubious originators issuing excess credit and inconsistent with these originators simply chasing prices, the ZIP codes with high concentrations of the worse originators exhibit a bust that is 63% larger than the bust of the matched ZIP codes.

Second, we use restrictive state antipredatory law changes (Bostic et al. 2008) introduced between 2004 and 2005 as a plausibly exogenous source of variation that restricts the lending activity of the worse originators. During the boom years, ZIP codes in states that passed antipredatory laws experienced a 9.6% lower home price increase annually relative to states with no law change. Third, we analyze prices around a pilot program in which risky loans were required to obtain mandatory examination and borrower counseling over a four-month period in the Chicago area as analyzed more generally by Agarwal et al. (2014). Though only a short-lived program, a sudden decrease in lending by the worse originators in treated ZIP codes is followed by a lower price appreciation as compared to the nearby similar control ZIP codes.

Fourth, another reverse causality possibility is that the worse originators were located in or entered urban markets with a tight land supply and that these originators focused on giving credit based on upward home price expectations. Elastic areas are not as subject to these concerns since there is limited upside due to relatively cheap land supply. Yet, elastic MSAs could experience considerable decreases in house prices if excess credit were given out to borrowers who could not repay. Consistent with the excess credit leading to

[3] Even the originators with the lower levels of second-lien misreporting exhibited some small amounts of second-lien misreporting, as well as misreporting along other dimensions. Measurement problems should reduce the power of our tests and understate the impact of dubious origination practices.

oversupply and an ultimate price collapse, elastic areas with a high presence of the worse originators exhibited large price decreases. In contrast, elastic areas with better originators exhibited only minor crashes.

All of our tests are consistent with the hypothesis that dubious origination activity causes house price distortions. While it may be possible to construct an alternative explanation for each one of the previous tests in isolation, it seems unlikely that there is a coherent alternative explanation consistent with the various tests.

We then turn to examine the channel through which the worse originators affected prices. First, dubious originators could extend credit to uncreditworthy borrowers by originating loans to borrowers with higher stated risk profiles. Second, they may have induced a "race-to-the-bottom" with better originators giving out riskier loans to compete against dubious originators. Third, these worse originators could be poorer at screening their applicants than other originators. Fourth, through the underreporting of applicant risk, they could be granting loans to applicants with risk profiles that are even worse than stated. We find evidence only for the first and fourth explanations. Our set of worse originators issues loans that have a much higher expected probability of delinquency at origination. The interest rates that the worse originators charged their borrowers were stronger predictors of future delinquency than the interest rates the better originators charged, suggesting that the worse originators were better at screening loan applicants than their counterparts. Finally, using propensity score-matched loans, we find that the originators who engaged in second-lien misreporting may have engaged in full-doc loan and income misreporting.[4] Thus, these originators who engaged in large amounts of second-lien misreporting had bad practices in the sense that they gave credit to borrowers with higher stated risk profiles, while simultaneously underreporting the true risk profiles of their borrowers.

Our findings also directly relate to the debate between Adelino, Schoar, and Severino (2015, 2016) and Mian and Sufi (2009, 2015). Adelino, Schoar, and Severino (2016) argue that the effects of housing price increases and credit expansion were not limited to low-income ZIP codes, and thus the expansion of excess subprime credit cannot explain the housing price expansion. We agree with Adelino, Schoar, and Severino (2016) that the housing price increases and credit growth were not limited to low-income ZIP codes. However, within each income demographic, the expansion of credit, as well as house price increases, are related to the supply channel through dubious origination practices. This contrasts with the main claim of Adelino, Schoar, and Severino (2016) that "cross-sectional distortions in the allocation of credit were not a key driver of the run-up in mortgage markets and the subsequent default crisis." Adelino, Schoar,

[4] The loans these originators reported as full documentation defaulted at a higher rate than those from other originators even after controlling for other loan attributes and a ZIP-code-level propensity score loan matching. These loans had missing debt-to-income information over 99% of the time as compared with 16% for the better originators.

and Severino (2015) argue that home buying within ZIP codes was facilitated by wealthier borrowers within the ZIP code. We find that the divergence between HMDA self-reported income and IRS income within a ZIP code is strongly related to our measure of dubious origination. This is consistent with Mian and Sufi's (2015) interpretation of the HMDA-IRS income difference as a likely measure of income misreporting.[5] Finally, Adelino, Schoar, and Severino (2015) claim that credit was not the source of house price increases, but credit came after house prices. However, across ZIP codes, we find that price peaks were preceded by the collapse of dubious credit. Overall, we find little evidence to support the claims of Adelino, Schoar, and Severino (2015, 2016) that "misreporting…simply does not explain the patterns we show" and, more generally, others that also argue that origination practices did not affect house prices (Foote, Gerardi, and Willen 2012) or that the crisis was not related to problems in securitization incentives (Gorton 2008, 2009).

Rather than securitization alone, our evidence supports the hypothesis that dubious origination practices facilitated through securitization had an economically large distorting effect on house prices. Our goal is not to examine all causes of the housing price bubble as surveyed by Mayer, Pence, and Sherlund (2009) or Levitin and Wachter (2012).[6] As strong as our findings are between dubious origination practices and house prices, it is important to note that our measure does not capture the full extent of misreporting. This should have the extent of dampening our findings, and thus, the true relation between dubious origination and house prices may be substantially stronger than that we document. We look forward to seeing more research more fully explore the relation between fraud and its economic costs.

1. Hypotheses

Housing prices respond to a shift in the demand curve (Herring and Wachter 1999; Hubbard and Mayer 2009). As lenders loosen credit standards, those who could not previously qualify to purchase a house are able to do so. Additionally, borrowers who qualified for smaller loans can afford larger ones. If lenders allow uncreditworthy purchasers to borrow large amounts of credit, then there could be a large shift in the demand for housing. The magnitude of the demand shift will depend on what fraction of new borrowers, who were previously credit-constrained, are given access to credit. Thus, our tests follow a rationale

[5] Adelino, Schoar, and Severino (2015) argue that misreporting cannot explain their results because their results hold in areas that they deem unlikely to have misreporting. However, their proxies for misreporting (based on GSE origination and subprime lenders) are weak proxies. In contrast to these claims, our proxy shows that the link between misreporting and both credit and house prices are extremely strong.

[6] Demyanyk and Van Hemert (2011) find deteriorating underwriting standards. Corbae and Quintin (2014) and Kermani (2012) argue that an increase in credit due to lower standards contributed to the boom of housing prices and to the subsequent bust.

similar to that of Mian and Sufi (2009) and Pavlov and Wachter (2011); when the supply of credit is increased by lowering underwriting standards, the demand curve for housing shifts outward, and housing prices increase.

Mortgage originators may affect demand differently. Originators who are willing to offer loans to uncreditworthy borrowers may shift the demand curve more than originators who screen borrowers to meet certain standards. Once an originator is willing to misreport whether a borrower has money down or income above a threshold, the loan may be issued to a borrower with little ability to repay.[7] This reasoning leads to our first hypothesis:

Hypothesis 1.1: ZIP codes with a larger fraction of originators with dubious practices experience larger house price increases during periods of credit expansion.

Hypothesis 1.2: ZIP codes with a larger fraction of originators with dubious practices experience larger price decreases during periods of credit contraction.

Alternatively, it may be that lending standards were low across the whole industry and that all originators gave out credit indiscriminately to uncreditworthy borrowers, mainly through nonagency securitization. If this is the case, then house prices should purely be related to the fraction of loans in the ZIP code that are securitized, independently of who is the originator. House prices should not be related to the market share of originators with misreporting in the ZIP code.

Hypotheses 1.1A and 1.2A: The misreporting practices of the originator (and therefore its presence in an area) have little or no relation to house price increases or decreases.

Saiz (2010) and Glaeser, Gyourko, and Saiz (2008) show that the elasticity of land supply has a large effect on housing prices. In inelastic ZIP codes, prices may increase quite rapidly along with increases in housing demand. On the contrary, Glaeser, Gyourko, and Saiz (2008) show that prices are never more than 10% to 15% above production costs in areas of elastic supply. In elastic areas, excess credit should not lead to larger price increases since new housing can be built as a response to the increased demand. However, it can lead to overbuilding, and when credit is removed, house prices may experience large busts. What is special about the hypotheses for dubious originators in elastic regions is that the excess credit may lead to large overbuilding of housing and a subsequent house price collapse even though there is little reason for a home price speculator to think that prices should fluctuate in this region. Thus, in

[7] This could occur for all types of securitized loans, not only for nonagency loans, but also for agency loans, to the extent that the misreporting was undetected by the government-sponsored enterprises (GSEs). In contrast, an originator who either is not securitizing or is securitizing loans, but not misreporting, may not lend to borrowers who are below standards. Additionally, Ben-David (2014) shows that higher leverage buyers paid an excess of 3.4% for the house, providing a more immediate channel for origination practices to affect house prices since misreported second-lien loans are typically of extremely high combined loan-to-value (LTV).

elastic areas, we can focus on the implications of potential excess lending and a price crash.

Hypothesis 2: In areas of elastic land supply, dubious origination practices lead to large price decreases when excess credit is no longer available.

Hypothesis 2A: In areas of elastic land supply, dubious origination practices have no relation to prices when excess credit is no longer available.

We now turn to our data and the measurement of dubious origination practices.

2. Data, Measures, and the Sample

In this section, we discuss our data sources, the measure of bad origination practices, the construction of our empirical measures, and the sample selection.

2.1 Data

The data used in this study are from a number of reputable sources. Property transaction information is obtained from DataQuick; securitized loan information is from Lewtan's ABSNet Loan; ZIP-code-level house price indices are from Zillow; ZIP-code-level demographics are from the Decennial Census 2000; ZIP-code-level household income information is from the IRS; and loan application information is from the Home Mortgage Disclosure Act (HMDA) data set.

DataQuick is one of the main providers of real estate transaction information recorded by county assessors. Specifically, we use DataQuick's History File, which provides the transfer date, location, the type of property transaction, and the names of the originators involved. Lewtan, on the other hand, compiles and cleans information from servicer/trustee reports of nonagency RMBS deals. We obtain house price indices from Zillow, an online real estate database. Mian and Sufi (2009) report that the Fiserv's Case Shiller Weiss indices have a correlation of 0.91 with overlapping ZIP codes in Zillow, yet Zillow's coverage is much broader.[8] Lastly, we obtain demographic and income data from the 2000 Decennial Census and the IRS SOI to use as controls in our empirical tests. We also use the HMDA data set, which contains detailed information about loan applications and the actions that followed the applications (i.e., whether the loan was originated and, in cases in which the origination failed, the reason why). Using census tract reference maps, we mapped the HMDA data to our ZIP codes for approximately 70% of our sample.

[8] Zillow covers 12,614 ZIP codes from 2003 to 2012. The Zillow Home Value Index is a time series of median home values. The Zillow median value is adjusted for seasonality and systematic residual error, among other filters. A detailed description of the methodology can be found at: www.zillow.com/research/zhvi-methodology-6032.

2.2 Originator practices

We use the measure of unreported second-lien loans in Griffin and Maturana (2016) as a proxy for mortgage originators' bad practices. This indicator essentially compares the information from servicer/trustee reports (in ABSNet) with the corresponding property transactions from the county deed records (in DataQuick). Although Griffin and Maturana examine three types of misreporting, they find that second-lien misreporting is directly attributable to mortgage originators, whereas owner occupancy misreporting and appraisal overstatements do not vary as much across originators. They find that more than 13% of the first-lien loans originated between 2002 and 2007 that were reported as not having a second lien in the RMBS records did have a second lien issued on the same day in the county-level transaction records. Piskorski, Seru, and Witkin (2015) also find extremely similar levels of second-lien misreporting using entirely different data sources and methodologies. They also find that misreporting varies widely across states, suggesting substantial cross-sectional variation for analysis. Griffin and Maturana find that the unreported second-lien indicator varies significantly across the set of the largest twenty-five mortgage originators in their sample. They show that delinquencies by the originator are strongly related to second-lien misreporting levels, even after controlling for the three types of loan-level misreporting. This suggests that the originators with high levels of second-lien misreporting may have engaged in other bad practices that led to losses.[9] In the last section, we further investigate potential income and documentation reporting issues with these originators. It is important to note that our approach is to classify the dubiousness in a ZIP code based on the originator rather than to calculate it directly. This approach is better if dubious originators engaged in other bad practices that are not directly measured, and the approach considers both agency and nonagency loans in a ZIP code. Nevertheless, we are not trying to distinguish the extent to which all of the increase in credit from originators who engaged in misreporting is due to fraud or simply lax standards by these originators.

Each year, we classify the same top twenty-five loan originators in Griffin and Maturana (2016) into three groups based on the cumulative fraction of misreported loans they issued. Specifically, we use the amount of cumulative second-lien misreporting of each originator from 2002 to year $t-1$ to rank the originators in year t.[10] We refer to the originators in the tercile with the highest misreporting as the *worse originators* and to the originators in the tercile with

[9] WMC Mortgage, which had the highest level of second-lien misreporting rate in Griffin and Maturana's sample, is reported to be under criminal investigation by the FBI and the U.S. Department of Justice. The accusations against WMC include rampant practices of falsifying loan documents in many dimensions and sidelining employees who repeatedly reported some of the falsifications they had seen (Hudson and Reckard 2012).

[10] Since we have the unreported second-lien indicator for the period 2002 to 2007, our rank of originators starts in 2003 (using only 2002 data). Also, beginning in 2008, we hold the ranking fixed for the following years. Table IA.1 shows the frequency with which each lender ranked in each tercile of second-lien misreporting between 2003 and 2008.

the lowest misreporting as the *better originators*.[11] Some of the originators with medium or lower levels of second-lien misreporting still have nonzero levels of second-lien misreporting (as shown in Figure IA.1). Additionally, some of these originators have been reported in the media to have engaged in questionable loan practices. For this reason, we call them "medium" or "better" but also note that they may have engaged in additional types of misreporting. To the extent that our benchmark for better practices still contains questionable origination practices, our empirical tests using this benchmark should understate the extent to which dubious origination practices affect house prices.

2.3 Empirical framework

We use DataQuick to create a set of measures to capture the importance of each type of originator in the mortgage market in each ZIP code. The use of the county deed records is important for expanding the set of mortgages to beyond those securitized in nonagency RMBS.[12] We use purchase transactions and not refinances since we are primarily interested in the transactions that influence market prices. To capture the relative importance of the different originator groups, we divide the number of loans for purchase each type of originator issued (worse, medium, better, and unranked) in each ZIP code by the total amount of purchase transactions in the ZIP code with originator information over the ranking period.

2.4 Sample selection

To increase the accuracy of our measures and empirical tests, we impose some restrictions on the ZIP codes to ensure adequate coverage and classification as described in the Internet Appendix B. This leaves a total of 5,176 ZIP codes.[13] The 5,176 ZIP code sample has lender name coverage for 42.5% of the observations. We only use these loans with lender name coverage to generate our measures.

Descriptive statistics for the ZIP-code-level measures and controls are shown in Table IA.2. On average, the worse originators were responsible for 5.6% of loan issuances between 2003 and 2006,[14] while the medium and better originators have 17.3% and 11.1%, respectively, of the market of loans with originator names reported in the 5,176 ZIP codes.

---

[11] We also refer to the originators in the second tercile as *medium originators*. If the originators are not in the set of twenty-five originators, they are called *unranked originators*.

[12] Internet Appendix A details which originators conducted business with GSEs. From the five loan originators that we consistently ranked in the worse category, we confirm that at least three were involved in GSE securitization.

[13] Mian and Sufi's (2009) sample is slightly over 3,000 ZIP codes.

[14] Figure IA.2 shows that *Worse originators' market share* (from 2003 to 2006) varies considerably across ZIP codes.

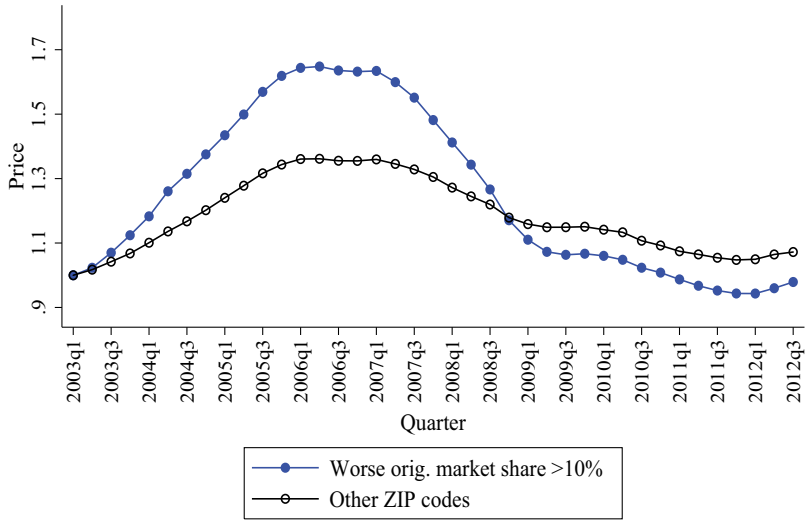

Figure 1
House price movements and worse originators' market share
This figure shows the relation between the activity of the worse originators and house prices. ZIP codes are divided into two groups: the first group contains ZIP codes for which the average market share of the worse originators during the period 2004Q3-2006Q2 exceeds 10% (solid circles), and the second group contains the remaining ZIP codes (hollow circles).

3. Origination Activity and House Prices

Our main goal in this section is to test whether house prices are related to the origination activity of misreporters (**Hypotheses 1.1 and 1.2**) or simply the amount of securitization (**Hypotheses 1.1A and 1.2A**).

3.1 General view

As a visual inspection, we first divide ZIP codes into two groups: the first group includes those in which the average market share of the worse originators during the third quarter of 2004 and the second quarter of 2006 exceeds 10%, and the second group is the remaining ZIP codes. These groups are composed of 858 and 4,318 ZIP codes, respectively.[15] Figure 1 shows the progression of house prices: the ZIP codes with the highest worse originators' market share went up by 63% during the 2003 to 2006 boom, whereas the ZIP codes with the lowest worse originators' market share only experienced a 36% run-up from 2003 to 2006. This 27% difference in absolute terms amounts to a 75% (27%/36%) relatively larger increase in house prices in the ZIP codes with the worse originators relative to the other ZIP codes. Conversely, from 2007

[15] Figure IA.3 shows that there is indeed a considerable difference in the average worse originators' market share during the 2003 to 2006 period between both groups. The worse originators' market share of both groups rapidly decreases towards zero during 2007, as most of the worse originators went bankrupt or lost considerable business.

to 2012, ZIP codes with the high presence of worse originators experienced a 40% decrease as compared to a 21% decrease for ZIP codes with a lower presence of worse originators, a 19% absolute difference or a 90% (19%/21%) higher relative decrease.

We look at this result from a geographic perspective (in Figure IA.4) and show that the 57% of ZIP codes in the top quartile of house price increases during the boom period and 61% of ZIP codes in the bottom quartile during the bust period have a high worse originators' market share. House price fluctuations and the worse originator market shares are particularly strong in the West Coast. While the correlation between misreporting and the regional patterns of house prices is quite interesting, we will exploit variation within MSAs in all of our main tests to see if our measure is capturing more than this simple correlation.

As additional motivation for our analysis, in Figure IA.5 we find that there exists a strong positive relation between nonagency securitization and ZIP code house returns. However, the simple univariate relation is eight to ten times stronger when using the fraction securitized by the worse originators rather than the total fraction of nonagency loans securitized.[16] This suggests that the proportion of dubious originators in a ZIP code explains the cross-sectional variation in ZIP code price movements much more than the total amount of nonagency securitization.

A central part of the debate between Adelino, Schoar, and Severino (2015, 2016), and Mian and Sufi (2015) centers upon the role of income. Mian and Sufi (2009) and Mayer and Pence (2009) show that house price distortions were concentrated in subprime ZIP codes. However, Adelino, Schoar, and Severino show that ZIP codes with high income also experienced substantial housing price increases, and hence they argue that the effect cannot be coming through subprime financing.

To place our results in the context of this debate, we independently sort ZIP codes by both their 2001 income and their level of misreporting originators over the period from 2003 to 2006. In panel A of Table 1, we report the average housing price return in the 2003 to 2006 run-up. Interestingly, there is little consistent relation between income quintiles and house prices, but there is an extremely strong relation between dubious origination market share and house price growth that holds in all five income quintiles, including the wealthiest. For example, for high-income ZIP codes, ZIP codes with low dubious practices experience a 26.3% price increase and ZIP codes with high dubious practices experience a 51.2% price increase. When looking at low-income ZIP codes, the ZIP codes with low dubious practices experience a 33.1% price increase and ZIP codes with high misreporting experience a 83% price increase. Hence, from ZIP codes with low to high dubious practices there is a 25% difference in home prices in high-income ZIP codes and a 50% difference for low-income

[16] Tables IA.3 and IA.4 confirm this result in a multivariate regression framework.

Table 1
House price returns by income and worse originators' market share

A. House price return, 2003–2006

	Worse originators' market share						
Avg. household income	Low	2	3	4	High	High - low	*t*-stat
Low	0.331	0.378	0.433	0.498	0.830	0.499	8.14
2	0.333	0.399	0.401	0.458	0.661	0.327	10.03
3	0.324	0.343	0.392	0.466	0.599	0.276	12.87
4	0.290	0.359	0.403	0.493	0.557	0.267	14.77
High	0.263	0.378	0.423	0.457	0.512	0.248	9.12
High - low	−0.068	0.000	−0.010	−0.041	−0.319		
t-stat	−2.44	0.00	−0.33	−1.23	−5.21		

B. House price return, 2007–2012

	Worse originators' market share						
Avg. household income	Low	2	3	4	High	High - low	*t*-stat
Low	−0.095	−0.152	−0.203	−0.305	−0.404	−0.309	−11.46
2	−0.105	−0.157	−0.202	−0.258	−0.348	−0.243	−14.38
3	−0.137	−0.182	−0.206	−0.241	−0.321	−0.184	−14.75
4	−0.145	−0.165	−0.211	−0.201	−0.290	−0.145	−13.37
High	−0.109	−0.141	−0.142	−0.170	−0.232	−0.123	−6.56
High - low	−0.014	0.010	0.061	0.135	0.172		
t-stat	−0.70	0.59	3.50	7.02	6.56		

This table shows the relation between house prices and income and the activity of the worse originators. ZIP codes are double sorted independently based on their average income in 2001 (as reported by the IRS) and *Worse originators' market share*. Panel A shows average house price returns from 2003 to 2006, and panel B shows average house price returns from 2007 to 2012.

households. Consistent with Mian and Sufi (2009), we find that the distortive effects of credit were most severe in low-income ZIP codes. Panel B of Table 1 shows that high-income ZIP codes with dubious origination have a 12.3% larger bust than ZIP codes with low dubious origination. Yet, in low-income ZIP codes, the difference in the decline in prices between ZIP codes with high and low dubious practices is 30.9%. This highlights that (1) there is an extremely strong relation between dubious origination and housing price returns and (2) the relation is considerably stronger in low-income ZIP codes.

There is one other feature of the table that should be noted. The sorts are independent, so the number of firms differs across bins. In Figure IA.6, we show the percentage of ZIP codes in each bin. Low-income ZIP codes had a much greater preponderance of misreporting originators than the high-income ZIP codes.

Overall, these results help to frame some of the debate in the literature. Mian and Sufi (2009) show that housing price fluctuations are concentrated in low-income ZIP codes. Our sorts confirm that this is where the majority of the fluctuations in housing prices lie. Adelino, Schoar, and Severino (2016) show that distortions in house prices are present in wealthy ZIP codes, which they interpret as evidence that the expansion in housing did not occur through excess credit channels. We show that indeed there are substantial fluctuations in high-income ZIP codes, yet the largest fluctuations are in wealthy ZIP codes

in which dubious origination holds a large market share. We will now explore these findings more thoroughly and later relate them to this debate.

3.2 Regression view

In panel A of Table 2, we present ordinary least-squares (OLS) regressions in which the ZIP code house price return is the dependent variable, and the market shares of the three different types of originators are the main explanatory variables of interest. The specification in Column 1 is without any controls, and in Column 2 MSA fixed effects are included (standard errors are heteroscedasticity robust and clustered by MSA to account for residuals being potentially correlated within economically connected areas). The worse originator market share in the ZIP code is strongly positively related to the ZIP code house prices, whereas the medium and better originators' market shares are weakly negatively related to ZIP code house prices. The relation between the three variables is not mechanical, as the unranked originators are excluded. We then include several controls. To account for the general relation between securitization and house returns, we include the total fraction of loans nonagency securitized at the ZIP code from 2003 to 2006. We also control for demographic characteristics that might be related to cross-sectional differences in house returns across ZIP codes, such as the ZIP codes' population, the number of house units, and the vacancy rate (all in the year 2000). Controls are also included for average household income in 2001 and average household income changes from 2001 to 2006 (also at the ZIP), as well as MSA fixed effects. Column 3 confirms the strong relation between the worse originators' market share and house returns during the boom. The coefficient of 1.235 on *Worse originators' market share* (which is statistically significant at the 1% level) implies that an increase of a 5% in market share by the worse originators in a ZIP code from 2003 to 2006 is associated with a house price increase of 6.18% on average during the boom.[17] Similarly, an increase of 5% in the worse originators' market share from 2003 to 2006 implies a decrease of 7.10% in house prices during the bust (Column 6). As in the boom, the relevance of the worse originators surpasses that of the other types of originators. Interestingly, the fraction of loans securitized does not enter significantly as a determinant of the housing price run-up during the boom with the inclusion of the worse originators' market share. The coefficient does enter significantly in the bust, though the slope is about one-seventh of that on the market share of the worse originators. Interestingly, average income changes in the boom are insignificant, though busts are less severe in ZIP codes in which income has increased.

[17] A 5% increase in market share by the worse originators is less than the difference between the median *Worse originators' market share* and its 90th percentile.

Table 2
Effect of worse originator activity on house returns

A. All ZIP codes

	2003–2006			2007–2012		
Worse originators' mkt. share	3.253***	1.743***	1.235***	−2.282***	−1.982***	−1.420***
	(29.34)	(4.47)	(2.91)	(−34.89)	(−5.10)	(−3.55)
Medium originators' mkt. share	1.297***	−0.449***	−0.320**	−0.007	0.010	−0.017
	(14.68)	(−3.27)	(−2.60)	(−0.13)	(0.09)	(−0.18)
Better originators' mkt. share	0.469***	−0.698*	−0.553*	−0.182***	0.782***	0.607**
	(4.36)	(−1.79)	(−1.67)	(−2.87)	(2.77)	(2.33)
Fraction securitized			0.045			−0.222***
			(0.40)			(−3.10)
Population			0.005***			−0.002***
			(3.15)			(−3.54)
Housing units			−0.011***			0.006***
			(−3.12)			(4.21)
Housing vacancy rate			0.658***			−0.167***
			(4.59)			(−4.02)
Average household income			−0.001***			0.000***
			(−2.67)			(5.07)
Change in avg. household income			0.001			0.000***
			(1.35)			(3.77)
Constant	−0.010	0.506***	0.479***	−0.066***	−0.192***	−0.195***
	(−0.73)	(10.07)	(11.90)	(−8.24)	(−4.43)	(−4.72)
MSA FE	n	y	y	n	y	y
SE clustered by MSA	n	y	y	n	y	y
Observations	5,176	5,176	5,176	5,176	5,176	5,176
Adj. R-squared	0.28	0.80	0.81	0.23	0.75	0.76
B. High-income ZIP codes						
Worse originators' mkt. share	1.930***	0.659**	0.850***	−1.474***	−1.495***	−1.420***
	(6.04)	(2.44)	(2.74)	(−7.07)	(−3.59)	(−3.75)
Medium originators' mkt. share	1.278***	−0.107	−0.075	0.146	0.171	0.159
	(8.36)	(−0.82)	(−0.57)	(1.46)	(1.30)	(1.28)
Better originators' mkt. share	−0.200	−0.242	−0.358**	−0.094	0.452***	0.351**
	(−1.13)	(−1.26)	(−2.42)	(−0.82)	(2.73)	(2.27)
Fraction securitized			−0.066			−0.014
			(−1.08)			(−0.24)
Population			0.001			−0.001
			(0.37)			(−0.94)
Housing units			−0.002			0.005**
			(−0.69)			(2.00)
Housing vacancy rate			0.311***			−0.179***
			(3.61)			(−3.47)
Change in avg. household income			0.000*			0.000***
			(1.78)			(3.02)
Constant	0.080***	0.389***	0.389***	−0.098***	−0.174***	−0.179***
	(3.67)	(10.40)	(10.39)	(−6.88)	(−6.47)	(−6.49)
MSA FE	n	y	y	n	y	y
SE clustered by MSA	n	y	y	n	y	y
Observations	1,035	1,035	1,035	1,035	1,035	1,035
Adj. R-squared	0.20	0.82	0.83	0.05	0.70	0.72

This table shows OLS estimates for regressions in which ZIP code price return is the dependent variable, on the ZIP-code-level market share for various types of originators from 2003 to 2006. The regressions include different combinations of demographic controls and MSA fixed effects. Columns 1 to 3 show the results for the boom period (2003-2006), and Columns 4 to 6 show the results for the bust period (2007-2012). Panel A shows the regression results for all ZIP codes, and panel B only includes the ZIP codes in the highest income quartile in 2001. *t*-statistics are presented in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Panel B of Table 2 shows the same specifications in panel A for the set of ZIP codes in the highest quartile based on household income during 2001. This relation is strong both statistically and economically, but in the boom, the slope

is slightly lower here than it is in the full sample. In the bust, the coefficient is similar to that in the full sample.

Overall, we find strong support for **Hypotheses 1.1 and 1.2** and little support for house prices being related to just securitization (**Hypotheses 1.1A and 1.2A**).

4. Did Dubious Origination Cause House Price Distortions?

Although we find a strong relation between house prices and the market share of the worse originators, the relation need not be causal, or causality might be reversed. For example, the worse originators may have followed more aggressive business strategies by entering ZIP codes with increasing house prices. A second and related possibility is that the worse originators may have targeted (or had a preference for) geographic areas with a tight land supply. Third, despite the use of various controls, there may still be an omitted variable driving both bad origination and house price movements. We take a variety of approaches to investigate if dubious origination practices cause house price distortions.

4.1 Were the worse originators simply chasing house returns?

The worse originators might have simply been chasing large house price returns to expand business and quickly entering booming ZIP codes or might have been good at targeting ZIP codes in which house prices were expected to increase. If this were the case, then if one matched ZIP codes with similar levels of house price appreciation during the boom but different levels of worse origination market share, one would expect a similar house price drop in the bust as the prices reverted to their nonbubble price. In contrast, if the worse originators were giving out undeserved credit unrelated to fundamentals that caused excess price run-up, then one would expect house prices in ZIP codes in which dubious originators had a high market share to drop to lower price levels than those in other ZIP codes with a similar run-up.

To examine these hypotheses, we take the 858 ZIP codes in which the worse originators had an average market share of more than 10% between the third quarter of 2004 and the second quarter of 2006, and match each of them to a ZIP code in the same MSA with the most similar housing returns from 2003 to 2006 in which the worse originators had a market share lower than 5%. The match within the same MSA indicates that geographic features, such as elasticity and economic fundamentals, will be similar though we will also control for ZIP code differences in the regression analysis below.

Panel A of Figure 2 compares the house price movements of the worse originator and the matched group. Consistent with the matching construction, the two groups on average have almost identical price run-up during the boom. For the worse misreporting group, the home prices decrease 39.4% on average, whereas for the ZIP codes with a lower presence of the bad originators matched

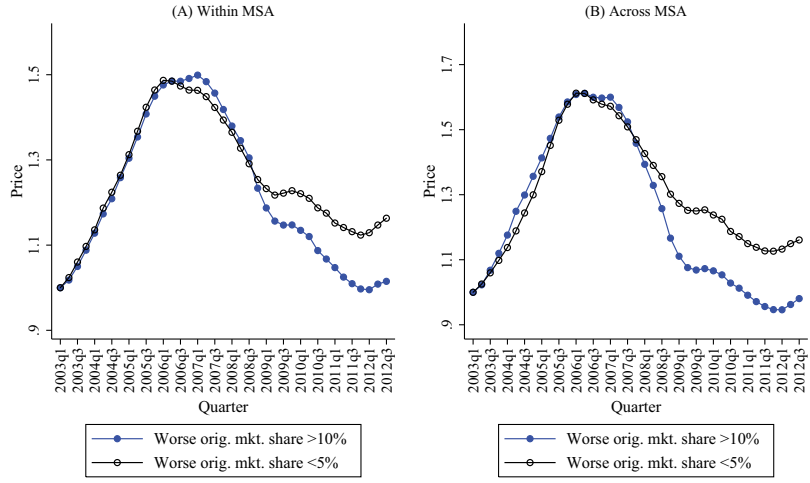

Figure 2
Run-up matching
This figure compares the average house price movement in ZIP codes for which the average market share of the worse originators during the period 2004Q3-2006Q2 exceeds 10% (solid circles) with the average house price movement in a group of ZIP codes that show an average market share of the worse originators below 5% during the same period (hollow circles). The control group is constructed to match the house returns of the group with high activity of the worse originators during the run-up period as closely as possible (matching is done with replacement, and ZIP codes are allowed to be matched a maximum of five times). ZIP codes in the control group are also required to be in the same MSA as the ZIP codes with high activity of the worse originators (*A*). Matching is done across MSAs (*B*). Note that when we construct confidence intervals from the cross-section of observations, the intervals are extremely small and close to the original lines. To avoid cluttering the graph, we do not include them.

within the same MSA, home prices decrease only by 23.5% on average. Thus, even though the two ZIP codes increase the same amount from 2003 to 2006, ZIP codes with bad originators experience a 15.9% larger drop in housing prices from 2007 to 2012.

In panel B of Figure 2, we repeat the previous exercise but with ZIP codes of different MSAs. The findings are similar; the ZIP codes with higher concentrations of the worse originators experienced the largest drops in house values after 2007.

One concern is that the differences in the bust could be due to differences in ZIP code characteristics, such as the average income between the two groups. Hence, in Table IA.5, we test the result shown in panel A of Figure 2 more formally by controlling for differences in the population, income, and growth in income in the ZIP codes. Column 2 in Table IA.5 shows that the house prices of the group of ZIP codes with more activity of the worse originators decreased an additional 14.6% on average compared to those of ZIP codes with less activity of the worse originators.

The previous effect is not explained by income, income growth, MSA, or other controls. The results are inconsistent with the notion that the relation between worse origination market share and home prices is due to the worse

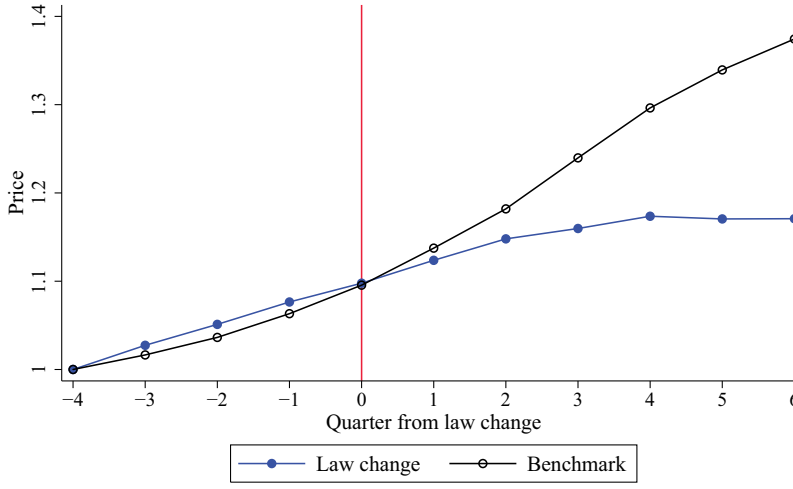

Figure 3
House price movements before and after APLs
This figure compares the house price movements in ZIP codes in states that passed restrictive antipredatory lending laws (APLs) between 2004 and 2005 (solid circles) with the house price movements in a benchmark of ZIP codes in states that did not pass any APLs before 2006 (hollow circles), before and after the law changes. The set of states that implemented restrictive APLs in 2004 and 2005 are New Mexico (Q1 of 2004), South Carolina (Q1 of 2004), Massachusetts (Q3 of 2004), Indiana (Q1 of 2005), and Wisconsin (Q1 of 2005). The set of states with no APLs are Arizona, Delaware, New Hampshire, Montana, Oregon, Washington, and Tennessee.

originators chasing house prices or targeting areas of increasing house prices. The results are consistent with the hypothesis that the worse originators issued unwarranted credit that caused a distortive effect on house prices.

4.2 Antipredatory law changes

First, we use antipredatory laws (APLs) as a quasi-natural experiment to analyze the effect of loan supply by the worse originators on house price movements. Bostic et al. (2008) find that APLs reduce subprime loan originations, especially when the APLs are more restrictive. These law changes should have also led to relatively fewer originations of loans by originators with worse standards. Hence, we compare house price movements in states that passed restrictive APLs between 2004 and 2005 with house price fluctuations in states with no antipredatory laws.[18]

Figure 3 shows the average house price movements of the ZIP codes that experienced a law change and of the ZIP codes in the benchmark (no APL).

[18] We confirm Bostic et al.'s (2008) relation between restrictive antipredatory law changes and subprime loan originations in Internet Appendix Figure IA.7. Also, as shown by Bostic et al. (2008), the set of states that implemented restrictive APLs in 2004 and 2005 are New Mexico (Q1 of 2004), South Carolina (Q1 of 2004), Massachusetts (Q3 of 2004), Indiana (Q1 of 2005), and Wisconsin (Q1 of 2005). The set of states with no APLs are Arizona, Delaware, New Hampshire, Montana, Oregon, Washington, and Tennessee.

Table 3
Effect of APLs on house price movements and worse originator loan supply

	All		High worse	
	ZIP codes		orig. supply	
	House returns	Supply	House returns	Supply
Postlaw	−0.024***	−0.021***	−0.035***	−0.051***
	(−4.31)	(−3.97)	(−7.03)	(−7.37)
Fraction securitized	0.035	0.221***	0.017	0.226***
	(1.72)	(6.05)	(1.55)	(6.58)
Population	0.000	0.001*	0.000	0.001
	(0.24)	(1.88)	(1.56)	(1.17)
Housing units	0.000	−0.002	−0.001*	−0.003
	(0.25)	(−1.62)	(−1.90)	(−1.24)
Housing vacancy rate	0.032*	0.043*	0.042**	−0.004
	(1.86)	(2.06)	(2.61)	(−0.16)
Average household income	−0.000	−0.000***	0.000	−0.000***
	(−1.55)	(−4.50)	(1.57)	(−5.92)
Constant	0.007	0.004	0.006	0.019**
	(1.17)	(0.96)	(0.84)	(2.61)
Quarter FE	y	y	y	y
Observations	17,162	17,000	8,710	8,880
Adj. R-squared	0.266	0.396	0.373	0.299

This table shows the effects of antipredatory lending laws on house price movements and loan supply by the worse originators, during the boom period (2003–2006). We construct a measure of worse originators' loan supply by dividing the number of loan originations by the worse originators each quarter by the total amount of loans granted by the worse originators from 2003 to 2009. To put the variable on a quarterly basis, values are then scaled by multiplying the variable by 28 (the number of quarters between 2003 and 2009). In the first two columns, the ZIP codes included are in states that passed antipredatory lending laws (APLs) between 2004 and 2005 or in states that did not pass any APLs before 2006. In the last two columns, the sample is restricted to the half of ZIP codes with the largest average loan supply by the worse originators. The variable *Postlaw* takes the value of one after the quarter in which an APL was passed, and zero otherwise. All regressions include quarter fixed effects. Reported *t*-statistics in parentheses are heteroscedasticity robust and clustered by CBSA. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Both sets of ZIP codes experience similar house price increases during the two-year period anteceding the law changes. However, after the law changes, house prices of ZIP codes in the first group continue to increase at a much slower rate than those of the ZIP codes in the benchmark.[19]

Table 3 shows the previous result more formally. We regress house price returns on a *Postlaw* dummy, a set of controls, and quarter fixed effects.[20] The negative coefficient on the law dummy variable of −0.024 (*t*-statistic of −4.31) means that ZIP codes in states that passed APLs had a 2.4% slower quarterly (9.6% annually) home price increase than in states with no APLs. The table also shows a negative effect of the law changes on worse originator loan supply (Column 2; coefficient of −2.1% and *t*-statistic of −3.97), which is consistent with the law change being the channel for reducing the dubious origination supply. Columns 3 and 4 of Table 3 show that the effects are

[19] In the Internet Appendix (Figure IA.8) we also plot the results for the three different quarters in which the law changes occur and find that average house prices significantly diverge in two of the three quarters (Q1 and Q3 of 2004).

[20] Standard errors are heteroscedasticity robust and clustered by core-based statistical area (CBSA), since clustering by a larger area, such as MSA, might yield insufficient clusters for the estimation.

considerably stronger for the subsample of ZIP codes with worse originator loan supply above the median level. The relative increase of house prices and worse originator loan supply of ZIP codes with law changes are 3.5% and 5.1% lower, respectively, than for ZIP codes with no APLs. We further validate the results in this section with a falsification test.[21] These findings are generally consistent with APLs preventing some bad quality loan originations that would have otherwise occurred and hence reducing upward pressure on house prices.

Second, we turn to a more localized but directed test. We exploit a pilot antipredatory legislative program (HB4050) that was implemented in ten Chicago ZIP codes from September 2006 to January 2007 in which riskier borrowers were required to undergo counseling and loan examination that included income verification. Agarwal et al. (2014) show that originators who specialized in riskier/subprime loans, fearing direct and indirect penalties for noncompliance, responded to the program by issuing fewer loans or directly ceasing operations in the pilot ZIP codes. Consistent with this result, panel A of Figure 4 shows the average monthly market share of the worse originators in the HB4050 ZIP codes experienced a substantially larger drop than in the benchmark ZIP codes.[22] More interestingly, panel B of Figure 4 shows the effect of the HB4050 program on house prices. From October 2006 to March 2007, house prices in the benchmark ZIP codes experienced a 60% relatively larger (3.5% as compared to 1.4%) house price increase on average than the HB4050 ZIP codes.[23]

In summary, both the law change and the HB4050 analyses are consistent with the worse originator activity distorting house prices during the boom (**Hypothesis 1.1**).

4.3 Are the price distortions by dubious originators explained by limited land supply?

We analyze the additional reverse causality possibility that worse originators were located in or entered urban markets with a tight land supply. In this scenario, the relation between home prices and worse origination activities would be related to the worse originators' geographical preference for certain areas associated with a higher probability of increasing home prices due to the inelastic land supply. Following Mian and Sufi (2009), we use elasticity

[21] Here, we falsely assume that the changes in the APLs occurred three quarters before the true date. Moving the APL occurrence three quarters allows New Mexico and South Carolina to have a prelaw period of two quarters. Table IA.6 shows that the false postlaw indicator has no effect on house prices and only a small effect on worse originator loan supply.

[22] The twelve ZIP codes in the benchmark identified by Agarwal et al. (2014) resemble the ZIP codes in the HB4050 area in terms of pretreatment socioeconomic characteristics and housing market conditions.

[23] Table IA.7 shows that the 2.1% house price differential is statistically significant through a difference-in-differences regression. Because of the small number of ZIP codes (22), we are unable to cluster standard errors by ZIP code, so we report heteroscedasticity robust t-statistics instead.

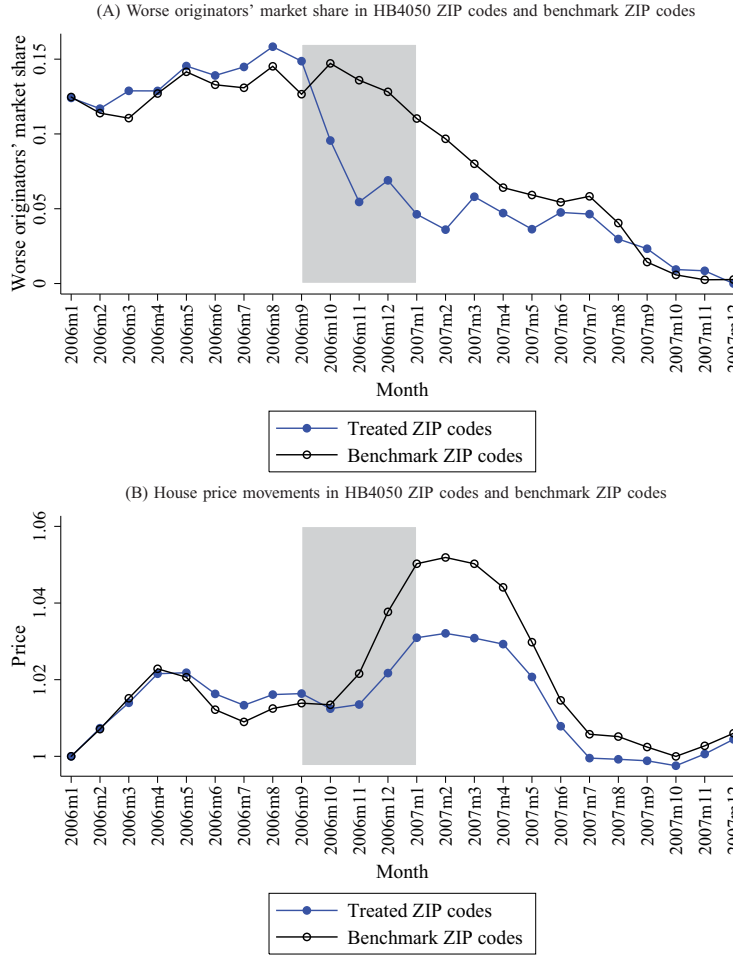

Figure 4
Worse originators' market share and house prices around the HB4050 program
This figure compares the market share of the worse originators and house prices in the ten ZIP codes in which the Illinois Predatory Lending Database Pilot Program (HB4050) was implemented, with the market share of the worse originators and house prices in the control group of twelve ZIP codes chosen in Agarwal et al. (2014). The ZIP codes in the control group resemble the ZIP codes in the HB4050 area in terms of pretreatment socioeconomic characteristics and housing market conditions (see Agarwal et al. 2014 for details). The gray area shows the period in which the pilot program was active. The figure shows the average monthly market share for both the treatment and the benchmark/control (*A*) and the average evolution of house prices in the two groups (*B*).

of housing supply from Saiz (2010) as a proxy for housing land supply.[24] The elasticity measure is a topologically based measure that gauges elasticity by

[24] Table VI of Saiz (2010) reports the elasticity of housing supply for the 1970 to 2000 period for ninety-five metro areas, each with a population over 500,000. We match sixty-five of these with our sample, which includes 90.2% of the ZIP codes.

Table 4
Effect of worse originator activity in elastic and inelastic ZIP codes

	Elastic MSAs		Inelastic MSAs	
	2007–2012		2007–2012	
	Top 50%	Top 25%	Bottom 50%	Bottom 25%
Worse originators' mkt. share	−1.809***	−2.400***	−1.268***	−1.265**
	(−5.73)	(−7.24)	(−2.99)	(−2.78)
Medium originators' mkt. share	−0.070	−0.161	0.076	0.042
	(−0.52)	(−0.88)	(0.56)	(0.23)
Better originators' mkt. share	−0.078	−0.544***	0.903***	1.127***
	(−0.26)	(−3.38)	(3.49)	(4.88)
Fraction securitized	−0.187	−0.177	−0.230***	−0.230**
	(−1.58)	(−1.61)	(−2.99)	(−2.58)
Population	−0.004***	0.000	−0.001***	−0.001**
	(−2.88)	(0.01)	(−2.98)	(−2.43)
Housing units	0.010***	−0.001	0.005***	0.005***
	(2.79)	(−0.21)	(4.01)	(3.33)
Housing vacancy rate	0.017	−0.094	−0.237***	−0.237***
	(0.15)	(−0.63)	(−4.55)	(−3.60)
Average household income	0.001	0.001	0.000***	0.000*
	(1.59)	(1.29)	(4.63)	(2.07)
Change in avg. household income	0.000	−0.000	0.000***	0.000
	(0.58)	(−0.52)	(2.90)	(1.27)
Constant	−0.062	0.061**	−0.292***	−0.336***
	(−1.66)	(2.29)	(−6.02)	(−6.35)
MSA FE	y	y	y	y
Observations	1,796	633	2,871	2,111
Adj. R-squared	0.67	0.67	0.76	0.70

This table shows OLS estimates for regressions in which ZIP code house price returns during the bust is the dependent variable, on the ZIP-code-level market share for various types of originators from 2003 to 2006, for different subsamples of ZIP codes based on housing supply elasticities from Saiz (2010). The regressions include different combinations of demographic controls and MSA fixed effects. Column 1 shows the estimates for the ZIP codes in MSAs in the most elastic half. Column 2 shows the regression for ZIP codes in MSAs in the most elastic quartile. Column 3 considers the most inelastic half, and Column 4 considers the most inelastic quartile. Reported *t*-statistics in parentheses are heteroscedasticity robust and clustered by MSA. ***p < 0.01, **p < 0.05, *p < 0.1.

surrounding geographic constraints. Glaeser, Gyourko, and Saiz (2008) show that house prices fluctuate considerably more in inelastic MSAs. Mayer and Pence (2009) show that subprime credit was in intercity areas, as well as in areas on the outskirts of cities. Yet the expansion of credit should have little effect on the run-up in prices in elastic MSAs, since increases in prices will predominately be due to increases in construction costs and limited increases in land supply. However, as discussed in **Hypothesis 2**, these areas could experience a considerable decrease in house prices during the crash if the increase in housing supply was fueled by an expansion of credit to unqualified borrowers that was not supported by income and population growth.

In Table 4, we estimate our main specifications for the bust for elastic and inelastic ZIP codes (for the top and bottom 50% and 25% of MSAs). For the elastic MSAs, a 5% increase in loan issuances by the worse originators explains an economically large decrease of 9.05% in house returns on average (Column 1). The results are slightly stronger in the top 25% of elastic ZIP codes (Column 2), consistent with the bust in inelastic ZIP codes being due

to bad origination practices. In Columns 3 and 4 (inelastic MSAs), the same coefficients are negative and significant as well, indicating that only the worse origination market share during the run-up (not medium or best) is associated with a bust in inelastic MSAs.[25] The fact that the origination activity by the worse originators during the boom is positively related to the bust in all levels of elasticity, and particularly in areas of elastic land supply, indicates that inelastic land supply is not the omitted variable driving the strong relation between dubious origination practices and house prices. In fact, although housing supply elasticity explains geographical differences in house price returns during the bust, *Worse originators' market share* has an effect more than twice as strong in the elastic ZIP codes. In addition, we regress house price changes on both the market share of the worse originators and housing supply elasticity, for which both variables are standardized. For each unit of standard deviation, the coefficient on the worse originator share has a 2.5 larger effect on house prices.[26]

As the worse originators expanded to very elastic areas, we further analyze the subset of ZIP codes in the 25% most elastic MSAs. Figure 5 shows that elastic ZIP codes with a low presence of the worse originators had only a minor burst in prices and ended up with house prices in 2012 around 20% above those in the ZIP codes with a high presence of the worse originators.

In summary, the fact that a high concentration of the worse originators is related to house price crashes in areas of elastic land supply indicates that the relation between dubious origination and crashes is not due to the worse originators solely concentrating in areas of tight land supply. The increase in credit in areas of elastic supply seemingly led to unwarranted housing construction and a subsequent crash of house prices.

While each test above may not accomplish identification in its purest form, it seems extremely difficult to construct a coherent alternative explanation that is consistent with all the previous results. Across all of our tests, the results are consistent with dubious origination practices causing house price distortions. We now seek to examine further the lending channel mechanism.

5. The Channel

We seek to learn more about the channel through which excess credit was granted. First, dubious originators may have lent to borrowers who had a higher ex ante delinquency rate. Second, better originators may have issued worse

[25] In Table IA.8 we present the same regressions for the boom. There is slightly less run-up in the highly elastic ZIP codes with more dubious originators. This might be due to a supply glut from dubious origination having effects in 2006. Most importantly, these run-up differences are considerably smaller than the large differences in the crash.

[26] A one-standard-deviation increase in *Worse originators' market share* implies a 5.7% drop in housing net worth from 2006 to 2009 on average, while one-standard-deviation change in housing supply elasticity has an effect of 2.3%, as shown in Table IA.9.

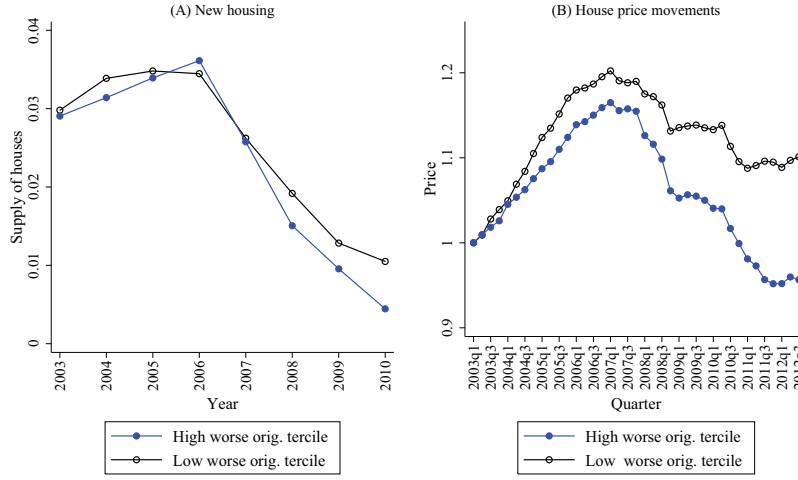

Figure 5
New houses and house price movements in elastic ZIP codes
This figure shows changes in new housing supply and house prices in elastic ZIP codes (ZIP codes in the upper quartile of more elastic MSAs). The figure shows the ZIP code average of new house transactions as a fraction of total houses in 2002 for the ZIP codes in the highest tercile of the worse originators' market share (solid circles) and the lowest tercile of the worse originators' market share (hollow circles) (*A*) and the average price changes for both groups (*B*).

quality loans in ZIP codes with a high presence of dubious originators. Third, dubious originators may have done a worse job of screening applicants, and this could explain their poor loan performance. Fourth, the process of understating borrower information may have caused uncreditworthy applicants to receive loans that they could not/would not repay.

5.1 Loan quality

We first estimate the credit risk of the loans at origination to see if the loans originated by dubious originators had similar or higher probabilities of delinquency. We then examine how ex ante loan risk varied for the better originators in ZIP codes in which the dubious originators had a high market share as compared to other ZIP codes. We base our estimates on the detailed stated loan and borrower characteristics in ABSNet for the nonagency market for the period before our main sample period.[27] We fit a logit model using all first-lien loans originated before 2001, where the dependent variable is a dummy that takes the value of one if the loan became seriously delinquent (90+ days) before 2002, and zero otherwise. This approach to estimate delinquency probabilities is similar to the one used by Ashcraft, Goldsmith-Pinkham, and

[27] A limitation of our analysis here is that it only use nonagency loans, which can be substantially different from agency loans (Keys, Seru, and Vig 2012). Jiang, Nelson, and Vytlacil (2014b) find that a large lending bank actually ends up holding some of the worse-performing loans on their books.

Vickery (2010). We then use the estimated coefficients in combination with the loan characteristics of the securitized loans originated during 2003 or later to obtain expected probabilities of delinquency.

Panel A of Figure 6 shows that in terms of the average ex ante probability of delinquency, the worse originators securitized loans with stated characteristics that were significantly worse than the loans securitized by the better originators. This occurs across ZIP codes in which the worse originators have both a high and low market share.

We examine if the competition with the worse originators might have led the better originators to issue riskier loans in ZIP codes in which there was a high presence of the worse originators. This does not appear to be the case. Quintiles 3 and 4 have higher probabilities of delinquency for the better originators, but the loans are actually relatively less risky in the ZIP codes with the highest presence of the worse originators. Originators with high misreporting are also not issuing riskier loans in the ZIP codes in which they have the largest presence.

In a similar manner, it is interesting to examine whether the different types of originators engaged in more second-lien misreporting in the areas in which the worse originators had a high market share. Panel B of Figure 6 shows lower levels of second-lien misreporting by the better originators that was not any higher in ZIP codes with a high market share of second-lien misreporters. It is important to note that misreporting is definitely nonnegligible for the "better" originators. This suggests that our results are an understatement of the effects of dubious practices. The worse originators do have higher levels of second-lien misreporting (50%) in the ZIP codes in which they have the highest presence; but they still misreport around 35% of their loans in the ZIP codes in which they do little business. This indicates that the misreporting practice was not primarily a problem only for certain loan officers or branch locations but a business practice and culture across the entire loan originating firm. This validates our approach of classifying ZIP codes based on originator composition rather than by simply captured misreporting, which can be measured less precisely in a ZIP code with fewer observations.

We also take a ZIP-code-level approach to the credit expansion. We examine whether "dubious" originators issued more loans in ZIP codes in which applicants previously could not receive credit because of their riskiness. We estimate OLS regressions in which the dependent variables are the market shares of the different types of originators between 2003 and 2006, and the independent variable of interest is the ZIP-code-level HMDA loan rejection rates (unmet demand) between 1996 and 1999. Of the three types of market share (i.e., worse, medium, and better), only the one corresponding to the worse originators is positively related to unmet demand (as shown in Table IA.10). This result is consistent with the worse originators expanding by granting credit to previously unmet demand from risky borrowers.

Overall, the fact that better originators are not engaged in riskier lending or higher levels of misreporting in ZIP codes with a high presence of originators

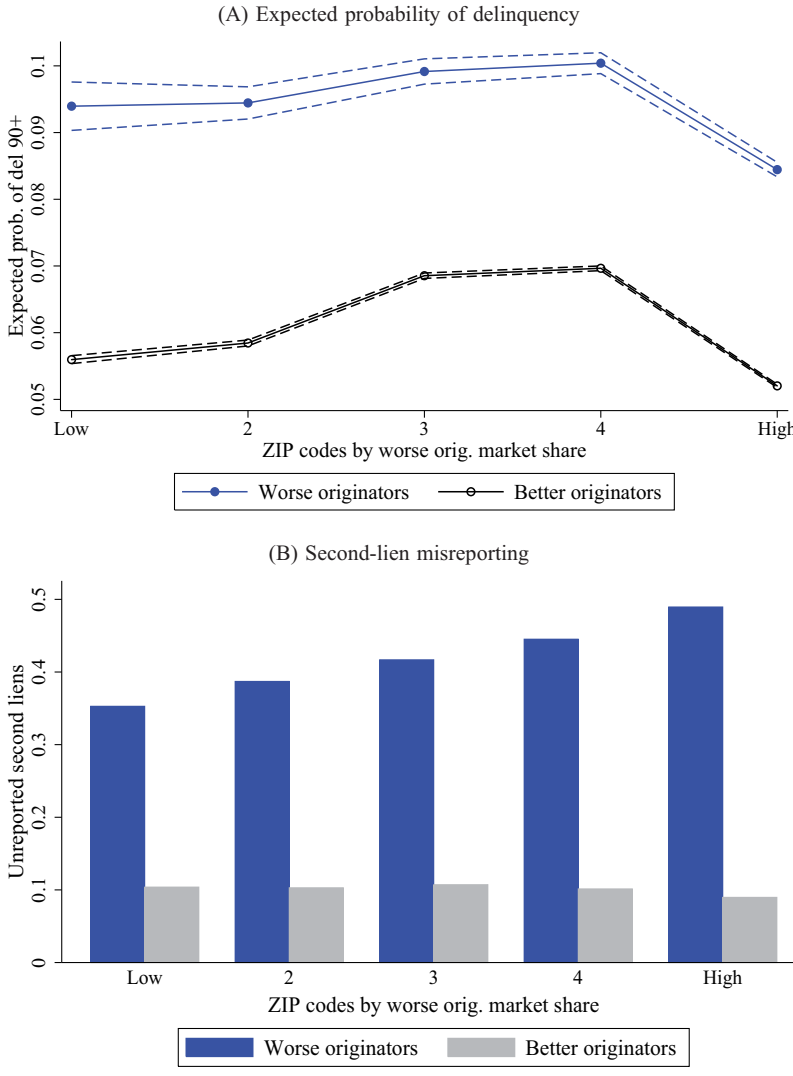

Figure 6
Worse and better originator quality comparison
This figure compares the quality of the loans originated by the worse and the better originators. The average expected probability of delinquency (90+ days) exhibited by loans issued by the worse and the better originators, by worse originators' market share quintile, is shown (*A*). The expected probability of delinquency is obtained by fitting a logit model at the beginning of 2002 using all first-lien loans originated before 2001 in ABSNet (this gives loans a period of at least a year to become delinquent). Specifically, the dependent variable is a dummy that takes the value of one if the loan became delinquent before 2002, and zero otherwise. The set of explanatory variables includes credit score, combined loan-to-value ratio, interest rate, the log of the loan amount, and dummy variables for level of documentation (low/no-doc or full-doc), self-reported occupancy status, refinance, and the existence of a prepayment penalty. The average of ZIP-code-level second-lien misreporting exhibited by purchase loans issued by the worse and the better originators, by worse originators' market share quintile, is shown (*B*).

with second-lien misreporting suggests that the effect of dubious origination in these ZIP codes is entering through the higher market share of the worse originators in the ZIP code. These worse originators tend to grant loans to borrowers with higher credit risk, some of whom may not have been able to receive credit previously.

5.2 Were misreporting originators poor at screening loans, or did they misreport loan quality?

We wish to understand if the worse originators were poor at screening borrowers, or if they understood that certain borrowers were of higher risk but lent to them anyway. If bad originators simply did a worse job of screening borrowers, the interest rate would be a less accurate predictor of delinquency for the loans by the worse originators than for the loans by the better originators. If an originator sought to maximize short-term profits, they might lend to a risky borrower at a high rate, but then underreport some of the loan's risky features when they resold it. In this case, the loan's interest rate would still be a good predictor of future delinquency, but the other loan characteristics reported at the time of securitization would be relatively weaker predictors. To investigate these possibilities, we estimate an OLS regression from 2003 to 2007 in which the dependent variable is an indicator of delinquency (90+ days), and the explanatory variables are a set of loan-level characteristics. We include the interaction of a dummy variable for worse originator with each of the variables.

In Column 1 of Table 5 we regress delinquency on the interest rate at the time of loan origination. Overall, the interest rate strongly forecasts delinquency, but it is a significantly more important predictor for the loans issued by the worse originators. In Column 2, with the other variables, we again find that interest rates are a much stronger predictor of delinquency for the loans issued by the worse originators. For this specification, we further explore this relation by examining the marginal R^2 in separate regressions for worse and better originators and report the results in Table IA.11. For better originators, interest rates lead to a relative increase in R^2 (i.e., delinquency forecasting) of only 0.2% as compared to 3.0% percent for the worse originators. Since the prices charged by worse originators were good determinants of future delinquencies, this suggests that the worse originators understood that some of their loans were riskier.

It is not clear if our separation of originators by second-lien misreporting is an issue related solely to second-lien misreporting or a symptom of other forms of misreporting. We obviously do not have access to internal bank sources, such as documentation and debt coverage, to verify such information. However, we can investigate the predictive value of loan attributes on delinquency. If an attribute was incorrectly reported, then this may decrease the variable's ability to predict future delinquency since the variable is reported with error. We find that the combined LTV ratio and the level of documentation are strong

Table 5
Explanatory power of loan-level controls

	All loans		Matched loans	
CLTV		0.609***		0.766***
		(33.52)		(31.67)
CLTV×Worse		−0.253***		−0.390***
		(−13.83)		(−14.82)
Full-doc		−8.538***		−9.912***
		(−20.30)		(−12.47)
Full-doc×Worse		1.400**		2.187***
		(2.32)		(3.77)
Interest rate	2.688***	0.851***	3.103***	1.024***
	(7.42)	(5.11)	(9.67)	(5.83)
Interest rate × Worse	1.445***	0.978***	1.066***	0.428***
	(15.52)	(9.47)	(11.33)	(2.83)
Nonowner occupied		2.900***		0.444
		(5.02)		(0.44)
Nonowner occupied × Worse		−2.736***		−0.600
		(−3.42)		(−0.76)
Credit score		−0.145***		−0.157***
		(−33.98)		(−18.05)
Credit score × Worse		0.003		0.011**
		(0.52)		(2.16)
ln(Loan amount)		4.128***		6.092***
		(9.79)		(7.71)
ln(Loan amount) × Worse		1.813***		2.631***
		(6.48)		(7.23)
ARM		1.391***		1.244**
		(5.51)		(2.00)
ARM×Worse		3.603***		3.341***
		(5.76)		(4.11)
Prepayment penalty		7.204***		10.539***
		(23.55)		(13.00)
Prepayment penalty×Worse		−2.516***		−4.912***
		(−4.50)		(−5.75)
ZIP×Year FE	y	y	y	y
Observations	932,236	932,236	173,644	173,644
Adj. R-squared	0.24	0.30	0.212	0.279

This table shows OLS loan-level regressions in which the dependent variable is an indicator for whether the loan became 90 days or more delinquent and the explanatory variables are a set of loan characteristics (from ABSNet). We also include the interaction of a dummy variable for worse originators with each one of the explanatory variables and ZIP code interacted with year of origination fixed effects. Columns 1 and 2 show the results for the full sample of loans. Columns 3 and 4 show the results for a sample where, for each loan issued by one of the worse originators, we find another loan issued by one of the better originators in the same ZIP code-year that also has similar propensity score. To compute the propensity score, we estimate a logit regression in which the dependent variable is a dummy that takes the value of one if the loan was issued by one of the worse originators and takes the value of zero if the loan was issued by one of the better originators. The explanatory variables are combined LTV, credit score, interest rate, the log of the loan amount, and indicators for low-doc, nonowner occupied property, arm loan, and the existence of a prepayment penalty. Also, we impose a maximum distance between propensity scores of 1%. We are able to impose such a tight criteria because there are many more loans from the better originators, and we match with replacement up to a maximum of five times. We find a match for 81% of the loans by the worse originators. Reported *t*-statistics in parentheses are heteroscedasticity robust and clustered by CBSA, and the regression's intercept is not reported. ***p < 0.01, **p < 0.05, *p < 0.1.

predictors of delinquency in general, but their explanatory power is significantly weaker for loan originators with high levels of second-lien misreporting. In regressions with and without CLTV ratios, CLTV ratios have three times as much forecasting power for better originators (Table IA.11).

Although these findings are consistent with potential misrepresentation of these loan features, it is also possible that the low predictive power of certain

borrower information is due to some other differences in the types of loans originated by the worse originators. The loans may have substantially different features that make the comparison of LTV ratio and documentation level problematic. To address this concern, we use a propensity score-matching approach in which for each loan issued by the worse originators, we find another loan issued by a better originator in the same ZIP code-year that also has similar combined LTV, credit score, interest rate, loan amount, and other characteristics.[28]

In Columns 3 and 4 of Table 5 we repeat the analysis discussed above in the matched sample of loans. We again find that combined LTV and the full-documentation indicator are significantly weaker predictors of delinquency for the worse originators, and this again raises the possibility that the worse originators further misreported along other dimensions. In addition, we estimate regressions of delinquency on loan characteristics for both matched groups (better and worse) separately in Table IA.13. Loan characteristics explain a much larger proportion of the variation in delinquencies in the sample of loans from the worse originators (difference in R^2 is 10% in absolute terms or 40% in relative terms).

The fact that the combined LTVs are weaker predictors of delinquency in the set of loans from the worse originators is to be expected given their higher levels of second-lien misrepresentation (as shown by Piskorski, Seru, and Witkin 2015 and Griffin and Maturana 2016). However, the fact that the same originators who engage in second-lien misreporting also have a lower predictive power of the full-documentation indicator raises the possibility of additional misreporting across this dimension.[29] To further investigate why full documentation leads to more delinquencies for the worse originators, we examine the reporting of debt-to-income for full-documentation loans.[30] Interestingly, worse and better originators have approximately the same percentage of loans self-reported as full docs (43.5% for the better and 43.1% for the worse). Yet of the loans classified as having full documentation in the ABSNet data, 16.9% of loans are missing debt-to-income for the better originators, but for the originators with high levels of second-lien misreporting, debt-to-income information is missing on 99.6% of the loans. These differences indicate either a lack of disclosure to trustees, which is more concentrated for the worse originators, or that the worse originators did not have proper

[28] We are able to impose a maximum propensity score difference of 1% and yet obtain 86,822 matches. Details of the matching process are provided in the header of Table 5, and the closeness of the matches is shown in Table IA.12.

[29] Indeed, as part of the settlement with the government, the JP Morgan statement of facts publicly admits to have identified "many instances because of missing documentation" in loans that were later securitized. A key piece of missing documentation was income. The report also identified excessive debt-to-income.

[30] Griffin and Maturana (2016) found that owner occupancy misreporting was primarily on behalf of occupants, and appraisal misreporting was primarily a misreporting from appraisers. Consistent with this finding, Figure IA.9 shows that owner occupancy misreporting and appraisal misreporting is similar between the better and worse originators.

documentation for some full-documentation loans. The fact that the full-documentation loans default at a relatively higher rate for dubious originators (as shown in Columns 3 and 4 of Table 5) lends support toward this second possibility. Jiang, Nelson, and Vytlacil (2014a) point to income falsification as a major issue, and Mian and Sufi (2015) show its importance at the ZIP code level.

In summary, we do not find evidence that the worse originators did a poor job of classifying risky borrowers since the interest rates they charged were actually better predictors of future delinquency than the interest rates charged by originators with lower levels of misreporting. This indicates that the worse originators knew that the borrowers were of high risk. We find that the originators who engaged in second-lien misreporting had considerably more full-documentation loans that both default at a higher rate and are missing a key piece of documentation (income). This raises the possibility that these originators engaged in other forms of misreporting beyond second-lien. These originators also expanded into ZIP codes with a high amount of previous loan rejections. Hence, we find that the "worse originators" were primarily "bad" in the sense of giving out credit to riskier borrowers and misreporting loan risk. These features are intuitively related, since an originator who can misreport key loan characteristics can potentially extend more credit to risky borrowers than an originator who correctly reports.

6. The Debate

Overall, our results support a loan supply-side channel explanation for the crisis. Our channel can be thought of as one similar to that originally proposed by Mian and Sufi (2009), but different in that the loan supply is primarily facilitated through dubious origination practices that are more severe in low income ZIP codes. In this section we focus on additional relations between our paper and three main aspects of the recent debate between Adelino, Schoar, and Severino (2015, 2016), and Mian and Sufi (2015).

6.1 Credit effect through loan supply?

A major point made by Adelino, Schoar, and Severino (2015, 2016) is that credit supply increased similarly in low and high-income ZIP codes. They argue that even in poorer areas credit mainly went to wealthier buyers, consistent with buyers' expectations of increasing house prices driving prices. Following their analysis, we now examine the number of new purchase loans, purchase loan size, and the total amount of new purchase credit over the period by looking at the same sorts in Table 1 on income quintiles and dubious origination. Panel A of Table 6 shows that poorer areas experienced more growth in the number

Table 6
Increase in loan supply from 2002 to 2006

A. Increase in number of loans

Avg. household income	Worse originators' market share							
	Low	2	3	4	High	Mean	High - low	*t*-stat
Low	0.083	0.086	0.098	0.124	0.140	0.106	0.057	2.40
2	0.065	0.071	0.086	0.111	0.100	0.087	0.035	2.35
3	0.038	0.057	0.084	0.103	0.074	0.071	0.036	2.77
4	0.038	0.055	0.092	0.037	0.026	0.050	−0.012	−1.24
High	0.018	0.043	0.030	0.016	0.060	0.033	0.042	2.53
Mean	0.049	0.062	0.078	0.078	0.080			
High - low	−0.065	−0.043	−0.068	−0.108	−0.080			
t-stat	−4.35	−2.93	−5.45	−7.95	−3.51			
B. Increase in average loan size								
Low	0.049	0.035	0.022	0.034	0.085	0.045	0.036	2.87
2	0.035	0.030	0.027	0.040	0.075	0.041	0.040	6.27
3	0.027	0.009	0.033	0.047	0.076	0.039	0.049	6.77
4	0.013	0.033	0.043	0.070	0.078	0.047	0.065	7.90
High	0.016	0.055	0.057	0.084	0.079	0.058	0.063	4.07
Mean	0.028	0.032	0.037	0.055	0.079			
High - low	−0.034	0.021	0.035	0.051	−0.006			
t-stat	−1.89	2.26	3.86	5.42	−0.55			
C. Increase in total loan amount								
Low	0.138	0.124	0.121	0.163	0.238	0.157	0.100	3.23
2	0.103	0.103	0.116	0.156	0.182	0.132	0.079	4.58
3	0.067	0.067	0.121	0.154	0.155	0.113	0.089	5.63
4	0.052	0.091	0.138	0.110	0.105	0.099	0.054	4.15
High	0.034	0.100	0.087	0.101	0.145	0.093	0.111	4.74
Mean	0.079	0.097	0.117	0.137	0.165			
High - low	−0.104	−0.023	−0.035	−0.062	−0.093			
t-stat	−4.46	−1.27	−2.37	−3.33	−3.19			
D. Increase in ZIP code income								
Low	0.024	0.023	0.023	0.023	0.025	0.024	0.001	0.27
2	0.026	0.028	0.026	0.026	0.026	0.027	0.000	−0.22
3	0.031	0.035	0.031	0.032	0.030	0.032	−0.002	−0.81
4	0.036	0.040	0.038	0.035	0.032	0.036	−0.004	−1.79
High	0.053	0.058	0.055	0.048	0.028	0.048	−0.025	−3.72
Mean	0.034	0.037	0.035	0.033	0.028			
High - low	0.029	0.034	0.032	0.025	0.003			
t-stat	4.23	8.41	10.75	10.23	0.75			

This table shows the relation between changes in loan supply and income and the activity of the worse originators. ZIP codes with more than 150 purchase originations per year in 2002 and 2006 are double sorted independently based on their average income in 2001 (as reported by the IRS) and *Worse originators' market share*. Panels A, B, C, and D show, from 2002 to 2006, the average annualized growth of the number of purchase originations, average purchase loan size, total purchase loan amount, and average income, respectively.

of new purchase loans from 2002 to 2006. Also, on average, the growth in the number of loans was higher in ZIP codes with more dubious origination.[31]

[31] We disagree with Adelino, Schoar, and Severino's (2015) assertion that "…only the relation between individual mortgage size and income that is informative about potential supply-side distortions." Unjustified new loans in an area may allow uncreditworthy borrowers to obtain mortgage financing, and this could be the marginal

Panel B of Table 6 shows that, on average, ZIP codes with a higher presence of dubious originators also experienced a larger increase in loan size. The average increase in loan size increases from 2.8% annually in the low worse market share bin to 7.9% in the bin with the highest market share of dubious originators.

Panel C of Table 6 focuses on the growth in the total amount of credit, which is the number of loans at their dollar amounts; low-income ZIP codes experience the largest percentage increase in credit, but the increase was considerably larger for those ZIP codes in which there was also a high presence of the dubious originators (annual increase of 23.8%). The table also shows that the increase in credit in high-income ZIP codes can be explained by the activities of the worse originators. For the ZIP codes in the highest quintile of income, the increase in credit supply was 14.5% per year in the ZIP codes with more dubious originators, while only a 3.4% per year in the ZIP codes with the lowest presence of the dubious originators.

Both the growth in the number of new purchase loans and the growth in average purchase loan size suggest that dubious originators seem to have been granting credit in a lax fashion in all the ZIP codes in which they operate. This is consistent with the excess credit channel and our tests above. Adelino, Schoar, and Severino (2015, 2016) argue that since ZIP codes in middle class and wealthy areas also have an expansion of credit, the credit expansion is unrelated to subprime financing and must be due to investor expectations. However, our findings that the expansion of credit is directly related to dubious practices in ZIP codes of various demographics is directly at odds with their arguments.

Finally, panel D of Table 6 shows that the ZIP codes with the largest increase in credit; that is, low-income ZIP codes with a high presence of the worse originators had no more income growth than other ZIP codes. Thus, the findings in this section are consistent with Mian and Sufi (2009) in the sense that low-income ZIP codes experience an increase in credit that was not justified by income growth and inconsistent with the arguments of Adelino, Schoar, and Severino (2015, 2016). There is a positive relation between income growth and dubious origination activity. Our tests indicate that after controlling for the income level, the increase in excess credit and prices was facilitated by dubious origination practices.

6.2 Income misreporting or cross-sectional differences in income?

Another key feature of the debate is whether the difference between HMDA and IRS income in a ZIP code is best thought of as a measure of income misreporting

borrowers that distort house prices. But, this disagreement is not central since we find that average loan size is also closely related to origination practices.

(Mian and Sufi 2015) or as a measure of the difference in the cross-sectional distribution of home buyers relative to the average person in the ZIP code (Adelino, Schoar, and Severino's 2015).[32] If *HMDA-IRS income* is related to misreporting practices, we would expect it to be related to our measure of misreporting originator activity. We find a correlation of 0.43 between *Worse originators' market share* and *HMDA-IRS income*, which is consistent with Mian and Sufi's income misreporting interpretation.[33] The tight relation we find is also consistent with our analysis above that dubious originators are more likely to engage in debt-to-income misreporting.

To further explore this relationship we classify ZIP codes based on their average household income in 2001 (as reported by the IRS) and worse originators' market share from 2003 to 2006 (as sorted previously in Table 1) and show the average difference between HMDA and IRS income (i.e., *HMDA-IRS income*) for each bin in panel A of Table 7. Interestingly, the average HMDA-IRS income difference increases with worse originators' market share for all household income bins.[34] Consistent with Mian and Sufi (2015), our findings show that misreporting is a more plausible explanation for the cross-sectional variation in HMDA relative to IRS income.[35]

Given that both *HMDA-IRS income* and *Worse originators' market share* seem to be correlated with dubious practices, it raises the question as to which of them has the strongest effect on housing prices. To answer this question, in panel B of Table 7 we present the results of regressions similar to those in Table 2 in which the ZIP code house price return is the dependent variable. Here, we standardize the explanatory variables for ease of comparison. The variable *Worse originators' market share* has the strongest effect on house price returns, both during the boom and the bust, in all specifications. More specifically, the effect of market share on house price returns is about 2.5 times stronger than that of *HMDA-IRS income* during the boom (Column 3) and about 24.5 times stronger during the bust.

6.3 The crash and loan supply

One problem with the investor expectations narrative for the crisis is that it is empirically vague about testable implications since there is no empirical

---

[32] We compute this measure, which we refer as *HMDA-IRS income*, and obtain an average value of 3.1% (with a standard deviation of 7.9%) for the 4,099 ZIP codes, for which we have both IRS and HMDA data.

[33] They also find a positive statistical relation between *HMDA-IRS income* and other fraud measures, though not as strong as the relationship with *Worse originators' market share*. The marginal effect of fraud on *HMDA-IRS income* ranged from 0.03 to 0.12 (see their Table 3). In particular, when we estimate similar regressions on *Worse originators' market share*, we find a coefficient of 0.72.

[34] We find similar results if we sort ZIP codes using HMDA income instead of IRS income, as reported in Table IA.14.

[35] It is worth noting that Adelino, Schoar, and Severino (2016) argue that misreporting should not affect HMDA income because its effect on credit supply is similar across subsamples of ZIP codes based on the fraction of GSE securitization and the presence of subprime lenders. Our combined results confirm that our market share measure is a much better and direct measure of bad practices.

Table 7
HMDA-IRS income and its effects on house price returns

A. HMDA-IRS income

Avg. household income	Worse originators' market share						
	Low	2	3	4	High	High - low	*t*-stat
Low	0.039	0.034	0.045	0.050	0.103	0.064	4.92
2	0.011	0.018	0.027	0.032	0.079	0.069	9.92
3	0.013	0.016	0.017	0.045	0.080	0.067	7.92
4	−0.005	−0.001	0.024	0.036	0.057	0.062	7.90
High	−0.030	−0.009	0.006	0.035	0.058	0.089	6.86
High - low	−0.070	−0.043	−0.039	−0.015	−0.045		
t-stat	−5.00	−4.34	−5.25	−1.75	−3.68		

B. Effect of worse originator activity and HMDA-IRS income on house returns

	2003–2006			2007–2012		
Worse originators' mkt. share	0.117***	0.067***	0.036*	−0.068***	−0.076***	−0.049**
	(26.14)	(3.82)	(1.96)	(−26.35)	(−4.32)	(−2.48)
HMDA-IRS income	0.071***	0.021***	0.014***	−0.028***	−0.010***	−0.002
	(15.76)	(3.31)	(4.05)	(−10.49)	(−4.11)	(−0.71)
Controls	n	n	y	n	n	y
MSA FE	n	y	y	n	y	y
SE clustered by MSA	n	y	y	n	y	y
Observations	4,099	4,099	4,099	4,099	4,099	4,099
Adj. R-squared	0.28	0.79	0.81	0.24	0.72	0.74

This table shows the relation between *HMDA-IRS income* and income and the activity of the worse originators (panel A), as well as the effect of worse originator activity and *HMDA-IRS income* on house price returns (panel B). In panel A, ZIP codes are double sorted independently based on their average income in 2001 (as reported by the IRS) and *Worse originators' market share*. Panel B shows OLS estimates for regressions in which ZIP code price return is the dependent variable on *Worse originators' market share* and *HMDA-IRS income*. The regressions include different combinations of demographic controls and MSA fixed effects. Columns 1 to 3 show the results for the boom period (2003-2006), and Columns 4 to 6 show the results for the bust period (2007-2012). *t*-statistics are presented in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

proposed measure of ZIP-code-level house price expectations. However, in the investor expectations narrative for the crisis, credit follows investor expectations and expectations should follow house price movements. Adelino, Schoar, and Severino (2015) states that households held larger mortgages because "they build on the availability of credit after house prices have appreciated: they are not the source of price increases." In contrast, a supply channel indicates that the credit is the source and causes price movements. An interesting feature of the crisis is that presumably because of poor loan originations, certain originators experienced financial difficulty when they needed to repurchase newly issued underperforming loans. This in turn presumably caused certain originators to cut back their supply of issuances. A supply-side explanation would argue that this lending should lead to a drop in house prices in the areas in which the credit is removed. The investor expectations explanation for the crisis posits that the credit from these originators would be driven by investor expectations and hence follow house prices.

These competing views can be tested as there is surprisingly considerable time variation in house price peaks between 2005 and 2007 (as shown in

Figure IA.10): 18.5% of house price peaks occur in 2005, 43.8% in 2006, and 31.3% in 2007.

Figure 7 aligns each of the dubious ZIP codes by the price peak and shows the quarterly ZIP code loan supply by the worse and better originators along with ZIP code house price movements during the four-year window around the house price peak.[36] Panel A consolidates across all peak-years. Both types of supply initially increase together with house prices and decrease before the house price peak. Loan supply by the worse originators decreases rapidly before the supply by the better originators. Panels B through D show that this pattern is consistent within each peak-year. In particular, each peak-year displays the pattern that supply by the worse originators peaked two to three quarters prior to prices. Note that for ZIP codes that peaked in 2005 and 2006, the supply of loans by the better originators is only slightly below its peak level six to eight quarters after the price peak even though prices in the ZIP have fallen 15%. These differences are confirmed by a test of difference in proportions.[37] Loan supply by the worse originators precedes price peaks more often than the loan supply by the better originators. The findings are consistent with a decrease of supply by the worse originators leading house price decreases (**Hypothesis 1.2**). The evidence is directly contrary to the assertion of Adelino, Schoar, and Severino (2015) that credit came "after house prices have increased." Moreover, the evidence also suggests that the worse originators may have played a major role causing the bust in house prices.

Overall, our results align with the securitization channel. We agree with Adelino, Schoar, and Severino (2015, 2016) that the effects of housing price increases and credit expansion were not limited to low-income ZIP codes. Yet, consistent with Mian and Sufi (2009), we find that the distortive effects of credit were most severe in low-income ZIP codes. In contrast to Adelino, Schoar, and Severino's claims and consistent with Mian and Sufi (2015), we find that the divergence between HMDA self-reported income and IRS income is generally related to misreporting, and not speculation from wealthy individuals in poor areas. Finally, across ZIP codes, price peaks were preceded by the collapse of dubious credit, and not by credit from dubious originators following house prices.

---

[36] Because we are focusing on the ZIP codes for which the excess credit was removed, we focus on the subset of ZIP codes for which the average market share of the worse originators exceeded 10% between the third quarter of 2004 and the second quarter of 2006.

[37] Table IA.15 shows that loan supply by the worse originators peaks before house prices in 90% of the ZIP codes (804 of the 858) for which house prices peaked between 2005 and 2007. Furthermore, this proportion is 24.3% larger in absolute terms (36.8% larger in relative terms) than the proportion of the better originators and is strongly significant (z-statistic of 11.7).

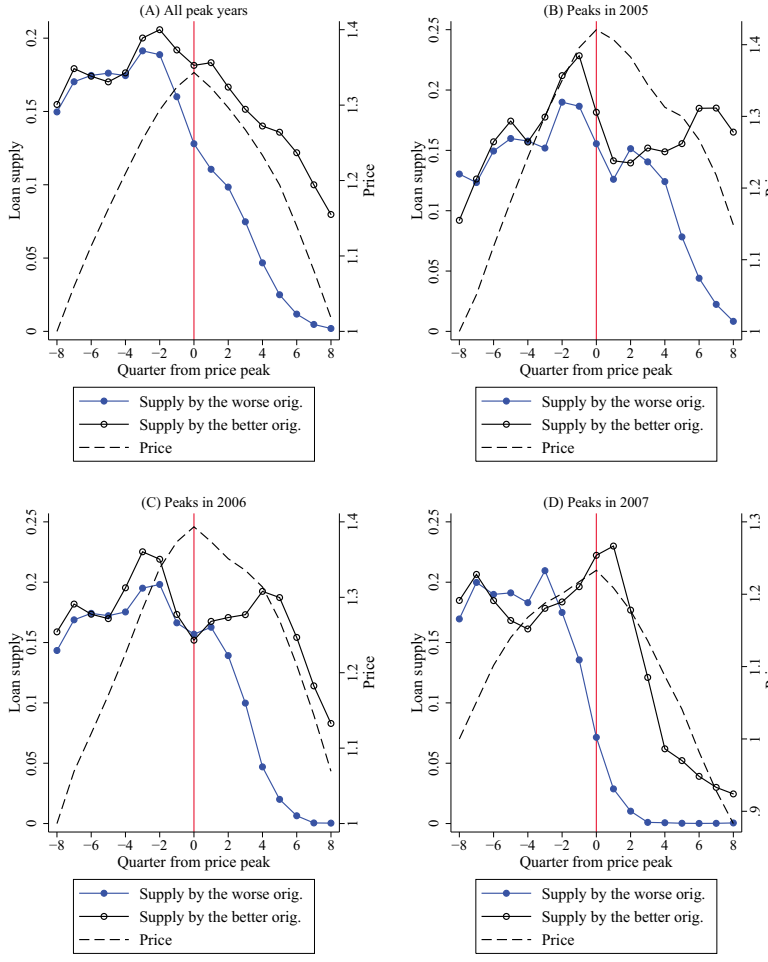

Figure 7
Loan supply and house price peaks
This figure shows the ZIP code-quarterly loan supply by the worse (solid circles) and the better (hollow circles) originators around ZIP code-house price peaks. The figure shows all ZIP codes (A) and ZIP codes for which house prices peaked in 2005, 2006, and 2007, respectively (*B*, *C*, and *D*). The dashed lines represent house price movements. Included are ZIP codes with an average market share of the worse originators during the period 2004Q3-2006Q2 exceeding 10%, where house prices peaked between 2005 and 2007.

7. Conclusion

The process of underreporting key loan attributes can have the by-product of facilitating credit to borrowers who have little ability to repay. We find that the presence of high concentrations of mortgage originators who engaged in second-lien misreporting in certain ZIP codes helps explain the 2003 to 2006 run-up of housing prices and its subsequent 2007 to 2012 collapse. The effect holds after controlling for total credit due to securitization, income, and income

growth and is present even in the wealthiest ZIP codes, indicating that it is not merely a subprime phenomenon.

We find no evidence to support the view that these effects are due to bad originators merely chasing prices or expanding into areas of highly inelastic house supply. The results are confirmed with a plausibly exogenous law change and a pilot lending examination program, indicating that causation is likely running from the lending activity of the dubious originators to house price distortions.

It is interesting to ask why the credit from these misreporting mortgage originators had such a large distorting influence on house prices. Since the interest rates that these bad originators charged were actually more useful predictors of delinquency than the interest rates charged by better originators, it seems these dubious originators were not worse in screening applicants. Nor do we find evidence that a high presence of bad origination is associated with more risky lending or misreporting by better originators. We find that the lenders who engaged in second-lien misreporting gave credit to borrowers with a considerably higher stated risk profile, while seemingly underreporting the true loan risk along other dimensions.

Overall, our results support a loan supply-side channel explanation for the crisis as proposed by Mian and Sufi (2009). We show that excess credit was not merely a subprime or securitization phenomena but one of dubious origination practices facilitated through securitization. These practices occurred across all income demographics but were more prevalent in low-income ZIP codes. Although we do not find support for the crisis narrative in which home prices were fueled primarily by responsible lending to credit-worthy investors with positive home price expectations, it seems plausible that home price increases caused by dubious lending fueled positive expectations for house prices and further feedback effects through speculative demand. This would be an interesting area for future work.

Our paper also highlights the unintended hidden costs that mortgage misreporting may have had on those who bought homes in areas with substantial misreporting. It is possible that the distortionary costs caused by the bad practices may be even more costly than the direct losses suffered by RMBS investors. These findings also suggest that actions of agents who facilitated misreporting jointly helped cause the real estate crisis and that these agents should not simply blame investor losses on market conditions. Our findings support the idea that misreporting loan information, a seemingly benign form of fraud, can have broad unintended consequences, even in the most open and transparent of markets. We hope our findings will spur additional debate and research on the role of trust and integrity in financial markets.

References

Adelino, M., A. Schoar, and F. Severino. 2015. Loan originations and defaults in the mortgage crisis: The role of the middle class. Working Paper, NBER.

———. 2016. Loan originations and defaults in the mortgage crisis: The role of the middle class. *Review of Financial Studies*.

Agarwal, S., G. Amromin, I. Ben-David, S. Chomsisengphet, and D. D. Evanoff. 2014. Predatory lending and the subprime crisis. *Journal of Financial Economics*, 113:29–52.

Akerlof, G. A., and P. M. Romer. 1993. Looting: The economic underworld of bankruptcy for profit. *Brookings Papers on Economic Activity*, 2:1–73.

Ashcraft, A., P. Goldsmith-Pinkham, and J. Vickery. 2010. MBS ratings and the mortgage credit boom. FRB of New York Staff Report.

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

———. 2014. High leverage and willingness to pay: Evidence from the residential housing market. Working Paper.

Bostic, R. W., K. C. Engel, P. A. McCoy, A. Pennington-Cross, and S. Wachter. 2008. State and local anti-predatory lending laws: The effect of legal enforcement mechanisms. *Journal of Economics and Business*, 60:47–66.

Carrillo, P. E. 2011. Testing for fraud in the residential mortgage market: How much did early-payment-defaults overpay for housing? *Journal of Real Estate Finance and Economics*, 1–29.

Corbae, D., and E. Quintin. 2014. Leverage and the foreclosure crisis. Working Paper.

Demyanyk, Y., and O. Van Hemert. 2011. Understanding the subprime mortgage crisis. *Review of Financial Studies*, 24:1848–80.

Foote, C. L., K. S. Gerardi, and P. S. Willen. 2012. Why did so many people make so many ex post bad decisions? The causes of the foreclosure crisis. Working Paper.

Garmaise, M. J. 2015. Borrower misreporting and loan performance. *Journal of Finance*, 70:449–84.

Glaeser, E. L., J. Gyourko, and A. Saiz. 2008. Housing supply and housing bubbles. *Journal of Urban Economics*, 64:198–217.

Gorton, G. B. 2008. The panic of 2007. Working Paper, National Bureau of Economic Research.

———. 2009. The subprime panic. *European Financial Management*, 15:10–46.

Griffin, J. M., and G. Maturana. 2016. Who facilitated misreporting in securitized loans? *Review of Financial Studies*, 29:384–419.

Herring, R. J., and S. Wachter. 1999. Real estate booms and banking busts: An international perspective. *Wharton Real Estate Review*.

Hubbard, R. G., and C. J. Mayer. 2009. The mortgage market meltdown and house prices. *BE Journal of Economic Analysis & Policy*, 9.

Hudson, M., and E. S. Reckard. 2012. GE lending unit said to be target of U.S. probe. *Los Angeles Times*, January 20.

Jiang, W., A. Nelson, and E. Vytlacil. 2014a. Liar's loan? Effects of origination channel and information falsification on mortgage delinquency. *Review of Economics and Statistics*, 96:1–18.

———. 2014b. Securitization and loan performance: A contrast of ex ante and ex post relations in the mortgage market. *Review of Financial Studies*, 27:454–83.

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

Keys, B. J., A. Seru, and V. Vig. 2012. Lender screening and the role of securitization: Evidence from prime and subprime mortgage markets. *Review of Financial Studies*, 25:2071–108.

Levitin, A. J., and S. Wachter. 2012. Explaining the housing bubble. *Georgetown Law Journal*, 100:1177–258.

Mayer, C., and K. Pence. 2009. Subprime mortgages: What, where, and to whom? *In Housing markets and the economy: Risk, regulation, and policy: Essays in honor of Karl E. Case.*, Eds. E. L. Glaeser and J. M. Quigley. Hoboken: Wiley.

Mayer, C., K. Pence, and S. M. Sherlund. 2009. The rise in mortgage defaults. *Journal of Economic Perspectives*, 23:27–50.

Mian, A., and A. Sufi. 2009. The consequences of mortgage credit expansion: Evidence from the US mortgage default crisis. *Quarterly Journal of Economics*, 124:1449–96.

———. 2015. Fraudulent income overstatement on mortgage applications during the credit expansion of 2002 to 2005. Working Paper, NBER.

Nadauld, T. D., and S. M. Sherlund. 2013. The impact of securitization on the expansion of subprime credit. *Journal of Financial Economics*, 107:454–76.

Pavlov, A., and S. Wachter. 2011. Subprime lending and real estate prices. *Real Estate Economics*, 39:1–17.

Piskorski, T., A. Seru, and J. Witkin. 2015. Asset quality misrepresentation by financial intermediaries: Evidence from RMBS market. *Journal of Finance*, 70:2635–78.

Saiz, A. 2010. The geographic determinants of housing supply. *Quarterly Journal of Economics*, 125:1253–96.

Zingales, L. 2015. Presidential address: Does finance benefit society? *Journal of Finance*, 70:1327–63.