The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXXII, NO. 3 • JUNE 2017

Forced Asset Sales and the Concentration of Outstanding Debt: Evidence from the Mortgage Market

GIOVANNI FAVARA and MARIASSUNTA GIANNETTI*

ABSTRACT

We provide evidence that lenders differ in their ex post incentives to internalize price-default externalities associated with the liquidation of collateralized debt. Using the mortgage market as a laboratory, we conjecture that lenders with a large share of outstanding mortgages on their balance sheets internalize the negative spillovers associated with the liquidation of defaulting mortgages and thus are less inclined to foreclose. We provide evidence consistent with our conjecture. Arguably as a consequence, zip codes with a higher concentration of outstanding mortgages experience smaller house prices declines. These results are not driven by unobservable zip code or lender characteristics.

FORCED SALES OF REAL AND FINANCIAL ASSETS fetch prices below their fundamental values (Shleifer and Vishny (1992)). When markets are illiquid, such sales generate price spillovers that reduce the value of similar assets held by other market participants. This may lead to further defaults and price-default spirals (Kiyotaki and Moore (1997), Gromb and Vayanos (2002), Brunnermeier and Pedersen (2009)). While a large literature documents the existence of fire sales and their spillover effects (Shleifer and Vishny (2011)), little attention

*Favara is with the Federal Reserve Board. Giannetti is with the Stockholm School of Economics, CEPR, and ECGI. We thank Bruno Bias (the Editor); Michael Roberts (the Coeditor); an anonymous Associate Editor; two anonymous referees; Manuel Adelino; Gene Amromin; Elliot Anenberg; Eric Engstrom; Julian Franks; Paul Kalem; Anastasia Kartasheva; Amir Kermani; Benjamin Keys; Elena Loutskina; Steven Ongena; Karen Pence; Anthony Pennington-Cross; Uday Rajan; Avri Ravid; Amit Seru; Steve Sharpe; Shane Sherlund; Greg Udell; Vikrant Vig; and seminar participants at the American Finance Association, Western Finance Association, NBER Financing Housing Capital Conference, London Business School Finance Summer Symposium, New York University Stern School of Business, University of British Columbia Sauder School of Business, Federal Reserve Bank of Chicago, Board of Governors of the Federal Reserve System, Boston University, Georgia State University, University of California Irvine, University of Lausanne and EPFL, University of Amsterdam, Goethe University in Frankfurt, IDC Annual Conference in Financial Economics Research in Herzliya, EFA in Lugano, FIRS Conference in Quebec City, SED Conference in Seoul, and the Summer Meeting of the Econometric Society at University of Southern California. We also thank Mihir Gandhi for outstanding research assistance. Giannetti acknowledges financial support from the Jan Wallander and Tom Hedelius Foundation and the Bank of Sweden Tercentenary Foundation. This paper represents the views of the authors and not those of the Federal Reserve System or its Board of Governors. The authors do not have any conflicts of interest, as identified in the Disclosure Policy.

DOI: 10.1111/jofi.12494

has been paid to the ex post incentives of market participants to avoid the negative spillovers associated with forced asset sales.

The purpose of this paper is to show that lenders with a high share of collateralized debt in their portfolios internalize the spillover effects of liquidation decisions on collateral values and are inclined to renegotiate their debt to avoid price-default spirals. Using differences in U.S. local housing markets (census tracts or zip codes) during the 2007 to 2010 housing crisis, we find evidence that such incentives are at work and are economically significant.

The U.S. housing crisis is an ideal laboratory for testing this conjecture for three reasons. First, mortgages, the standard debt contracts in the housing market, entitle lenders to seize the houses and sell them through a foreclosure process if borrowers default. Second, as the housing market is illiquid, foreclosures may generate price discounts that tend to spill over to nondistressed neighboring houses (Harding, Rosenblatt, and Yao (2009), Campbell, Giglio, and Pathak (2011), Anenberg and Kung (2014), Hartley (2014)). Third, the recent crisis has seen an unprecedented increase in foreclosures and decline in house prices, with feedback loops between foreclosures and prices contributing to the severity of the crisis. For instance, it has been shown that foreclosures led to a generalized decline in house prices (Mian, Sufi, and Trebbi (2015)), which in turn caused additional foreclosures as borrowers moved into negative equity positions (Elul et al. (2010)), triggering further price declines (Guren and McQuade, 2013).

We begin the analysis with a stylized model of the housing market in which negative income shocks force distressed homeowners to default on their mortgage obligations, and foreclosures trigger a decline in house prices as liquidations create an imbalance of housing demand and supply. When mortgages are held by many (atomistic) lenders, each lender places no weight on the effects of its foreclosure decisions on local house prices, and therefore foreclosures are likely to be followed by further foreclosures. In contrast, when lenders hold a large share of the mortgages outstanding in a neighborhood on their balance sheets, they internalize the adverse effects of their liquidation decisions on house prices in that neighborhood and have stronger incentives to avoid foreclosures, mitigating the adverse effects of foreclosures on house prices.

To test this theoretical prediction, we perform two sets of tests. First, for a subset of lenders with available data on mortgage performance, we use loan-level data to test whether a lender's incentives to foreclose on defaulting mortgages depend on the proportion of the outstanding mortgages it has retained on its balance sheet in a census tract. Second, we use aggregate zip code-level data to test whether lenders' incentives have implications for house prices.

In the loan-level analysis, we measure the share of outstanding mortgages kept on lenders' balance sheets at the census tract level. We expect this measure to capture lenders' incentives to internalize foreclosure externalities in a neighborhood. Consistent with the predictions of our stylized model, we find that, while this share is negatively correlated with lenders' propensity to foreclose, there is no such correlation for securitized mortgages. We obtain these results in specifications that include lender and local market fixed effects to purge unobserved lender and local market characteristics that might correlate with the share of mortgages retained by the lender.

To mitigate the concern that lenders retain more mortgages in census tracts in which they have an informational advantage, all of our tests are conducted conditional on mortgages being 90^{+} days delinquent. The very fact that we focus on seriously delinquent mortgages should reduce concerns that lenders had a prior informational advantage on the quality of these mortgages. In addition, we exploit exogenous variation induced by mergers of nonfailing large banks to instrument the share of mortgages retained in a neighborhood. The nature and size of these mergers make it unlikely that they are related to local market characteristics. Our results are qualitatively robust and quantitatively stronger when we use this instrumentation strategy.

We also provide evidence consistent with other predictions of the model. For instance, the model implies that lenders' foreclosure decisions are strategic substitutes. This means that the likelihood that a lender forecloses on a defaulting mortgage depends negatively on the share of outstanding mortgages on its own balance sheet and positively on the share of mortgages retained by other lenders. The intuition is that house prices drop less when other lenders renegotiate defaulting mortgages, strengthening a lender's incentives to foreclose in order to maximize its payoff of liquidation. Our model also implies that lenders should be more inclined to foreclose in areas in which it is easier to resell foreclosed properties and less inclined to foreclose in highly distressed areas. In line with these predictions, we find that the share of outstanding mortgages retained by a lender reduces the propensity to foreclose to a lesser extent in areas with a higher share of local mortgages retained by other lenders and in highly desirable neighborhoods, and to a greater extent in areas with more distressed households. We find no evidence supporting alternative mechanisms, including differences in individual lenders' organizational capabilities to foreclose or renegotiate mortgages in different geographical areas.[1]

Next, we consider the aggregate implications of our hypothesis using zip code–level data. We perform tests at the zip code level because zip codes are the smallest geographical areas for which we can obtain comprehensive data on house prices, and arguably the largest areas within which foreclosures are likely to generate negative externalities on house prices.[2] We construct a zip code–level measure of the concentration of mortgages on lenders' balance sheets using data on mortgages retained by the four largest holders in a zip code. Zip codes with a higher concentration of outstanding mortgages are expected to experience lower foreclosure rates. We find empirical support for this conjecture even after instrumenting the concentration of mortgages on lenders' balance sheets with exogenous bank mergers. A one-standard-deviation change in the

[1] The finding that *within*-lender differences in organizational capabilities do not drive our results is not at odds with the findings of Agarwal et al. (2015) that loss mitigation in mortgage markets is explained largely by differences in renegotiation capabilities *across* lenders.

[2] See Mian, Sufi, and Trebbi (2015) for empirical evidence on foreclosure externalities at the zip code level.

concentration index leads to an approximately 3% lower rate of house price declines in a zip code.

Finally, to provide additional evidence consistent with a causal mechanism running from the concentration of outstanding mortgages to foreclosure rates and house prices, we study how the results vary across jurisdictions with different foreclosure procedures. We expect that any lender, regardless of its share of outstanding mortgages in the neighborhood, has weaker incentives to foreclose in states with costly foreclosure procedures. Consistent with this idea, we find that the concentration of mortgages on lenders' balance sheets is associated with fewer foreclosures only in nonjudicial states where foreclosure costs are lower. These results strengthen the interpretation that lenders differ in their ex post incentives to resolve distress and not in their ex ante ability to screen borrowers.

Our paper is most closely related to empirical research on forced sales of real and financial assets (Shleifer and Vishny (1992, 2011)). Part of this literature focuses on the negative externalities associated with asset sales in economic downturns (e.g., Asquith, Gertner, and Scharfstein (1994), Benmelech and Bergman (2011)). We depart from this literature by studying how lenders' incentives to avoid externalities due to liquidation depend on their credit exposure. To our knowledge, this is the first paper to explore the role of these incentives in lenders' liquidation decisions and asset prices.

As our analysis focuses on the housing market, our paper also contributes to the literature on the recent U.S. housing crisis. A number of papers explore how differences in local mortgage markets relate to the intensity of the crisis (e.g., Mayer, Pence, and Sherlund (2009), Mian and Sufi (2009, Mian and Sufi 2011)) and whether securitization has exacerbated the intensity of the crisis by inhibiting the renegotiation of delinquent mortgages (Piskorski, Seru, and Vig (2010), Adelino, Gerardi, and Willen (2013), Agarwal et al. (2015)). Our paper focuses, instead, on the incentives of lenders to foreclose portfolio mortgages, relates such incentives to the share of outstanding mortgages that they retained in a neighborhood, and studies the implications for foreclosure rates and house prices.

The paper is also related to the literature that explores the effects of banks' loan concentration on bank-firm relationships (Berger et al. (2005)), the loan supply (Garmaise and Moskowitz (2006)), and the transmission of monetary policy to mortgage rates (Scharfstein and Sunderam (2015)). All of these papers study the effects of market power concentration on loan origination and contract terms. In contrast, we focus on the role of the concentration of outstanding mortgages on lenders' ex post incentives, a mechanism unrelated to market power. By showing that a market with dispersed lenders is more prone to liquidation externalities, we also provide an alternative interpretation to the view that competition in the credit market erodes financial stability because it distorts lenders' risk-taking decisions by lowering their profit margins (Keeley (1990)).

The paper is organized as follows. In Section I, we develop a simple model that guides our empirical analysis. Section II describes the data. Sections III and IV present the loan level tests and the zip code–level tests, respectively. Section V concludes.

I. Theory and Testable Implications

In this section, we develop a simple model to illustrate the relationship between foreclosures, house prices, and the concentration of outstanding mortgages on lenders' balance sheets. In the model, foreclosures generate an imbalance of housing supply and demand and cause a decline in equilibrium prices. Lenders' payoffs depend on house prices because the liquidation value of the housing collateral falls as foreclosures increase.[3]

Since lenders' payoffs depend on house prices, interesting feedbacks arise between lenders' incentives to foreclose and house prices. When lenders hold a large share of outstanding mortgages on their balance sheets, they are more likely to internalize the potentially adverse effects of foreclosures on house prices, which reduces their propensity to foreclose. In contrast, lenders that have retained a small share of outstanding mortgages place little weight on the effects of their foreclosure decisions on house prices and always foreclose. In the following sections, we bring these model predictions to the data.

A. The Model

A.1. Assumptions

There are two dates and two groups of agents of mass one, households (indexed by k) and lenders (indexed by i). At $t = 0$, some households enter the period with one unit of housing endowment, $h_{0k} = 1$, and an outstanding mortgage payment, B. At $t = 1$, households enjoy utility from consumption, $c_k \geq 0$, and housing, $h_k \in \{0, 1\}$:

$$U_k = c_k + \gamma_k h_k,$$

where γ_k is uniformly distributed, $\gamma_k \sim \mathcal{U}[0, \overline{\gamma}]$, and captures heterogeneity in utility from homeownership. Households with endowment $h_{0k} = 1$ have the highest utility from housing services. Aggregate housing supply is fixed at $\overline{H} < \overline{\gamma}$.

At $t = 1$, there are two states of the world. In the bad state of the world, a fraction e of households suffers a negative income shock and receives zero income, while the remaining fraction receives w. In the good state of the world, all households receive w. We assume that income shocks are observable and that only part of households' income is pledgeable, that is, $B < w$.

Under these assumptions, households with a high-income realization always repay B. For households with zero income, lenders may partially recover B

[3] The assumption that lenders' payoffs depend on house prices may also capture the negative externalities of foreclosures, for instance, if the decline in house prices triggers more foreclosures as other borrowers move into negative equity positions. Lower house prices may also directly impair the balance sheets of lenders that hold previously repossessed properties.

by foreclosing their homes with probability λ and selling the houses at price p.[4] Below, we determine the equilibrium house price and lenders' foreclosure probabilities.

A.2. Equilibrium House Prices with Atomistic Lenders

In the absence of income shocks, the unit housing demand for household k is pinned down by the condition

$$\gamma_k \geq p,$$

which relates the utility value of owning to the price of housing. The equilibrium price is determined by equating aggregate demand and supply. Since γ_k is uniformly distributed, the equilibrium condition is

$$p = \overline{\gamma} - \overline{H}.$$

All households repay B and, under our assumption on the initial distribution of housing, hold onto their houses.

In contrast, when a fraction e of households is hit by a negative income shock, they cannot afford to repay B. If lenders foreclose on these households with probability λ, a fraction λe of them is excluded from the housing market and their homes $\lambda e\overline{H}$ will be on the market. The demand for housing of individuals that did not own a house at $t = 0$ and did not suffer a negative income shock at $t = 1$ is $(1 - e)(\overline{\gamma} - \overline{H} - p^L)$. Therefore, the market-clearing condition is

$$(1 - e)(\overline{\gamma} - \overline{H} - p^L) = \lambda e\overline{H},$$

and the equilibrium house price is

$$p^L = \overline{\gamma} - \overline{H} - \frac{\lambda e\overline{H}}{1 - e}.$$

It follows that p^L is strictly lower than p and decreases in the probability of foreclosure λ, because a fraction e of households with high utility from owning cannot participate in the market, reducing aggregate demand.

Any lender chooses λ to maximize its expected payoff:

$$(1 - e)B + e\lambda p^L.$$

It follows that atomistic lenders (that do not internalize the effect of λ on p^L) always find it optimal to foreclose with probability one in order to maximize the liquidation value of the foreclosed property. This finding is summarized in the following lemma.

LEMMA 1: *Atomistic lenders foreclose on distressed households with probability one. The equilibrium price of housing is* $p^L = \overline{\gamma} - \frac{\overline{H}}{1-e}$.

[4] Households will default instead of selling the house and repaying the debt if B is sufficiently large. To ensure this, we hereafter assume that $B = \overline{\gamma} - \overline{H}$.

This result provides a novel explanation for why securitization leads to more foreclosures.[5] Investors in securitized mortgages can be thought of as atomistic lenders, and should therefore have stronger incentives to foreclose defaulting mortgages. Importantly, these foreclosure decisions may cause inefficiencies because households that value houses the most become unable to keep their properties.

B. Lenders' Shares of Outstanding Mortgages and Foreclosure Decisions

Suppose now that one lender holds a large share, ξ_1, of the outstanding mortgages (the "large" lender), while the remaining share $(1-\xi_1)$ of mortgages is dispersed among many atomistic lenders. The large lender solves the same problem as the atomistic lenders but internalizes the effect of its foreclosure decisions on house prices. It may thus find it optimal to foreclose with probability $\lambda_1 < 1$ even though atomistic lenders continue to foreclose with probability one.

To see whether this is the case in equilibrium, assume that the large lender forecloses with probability λ_1, while atomistic lenders optimally foreclose with probability one. In this case, the market-clearing condition at $t = 1$ is

$$(1-e)(\overline{\gamma}-\overline{H}-p^{L'}) = \xi_1\lambda_1 e\overline{H} + (1-\xi_1)e\overline{H}.$$

The right-hand side is the demand for housing by the fraction $(1-e)$ of unaffected households. The left-hand side is the supply of houses that are put on the market when a fraction e of households is hit by a negative income shock, the large lender forecloses with probability λ_1, and atomistic lenders foreclose with probability one. The equilibrium house price is now

$$p^{L'} = \overline{\gamma} - \frac{\overline{H}}{1-e} + \frac{(1-\lambda_1)\xi_1 e\overline{H}}{1-e} > p^L.$$

The large lender maximizes the same payoff function of an atomistic lender, but internalizes the effect of its foreclosure policy on housing prices. Thus, maximizing $(1-e)B + e\lambda_1 p^{L'}$ is equivalent to solving the following problem:

$$\max_{\lambda_1}(1-e)B + e\lambda_1\left(\overline{\gamma} - \frac{\overline{H}}{1-e} + \frac{(1-\lambda_1)\xi_1 e\overline{H}}{1-e}\right).$$

The optimal foreclosure probability is

$$\lambda_1 = \min\left\{\frac{1}{2}\left(1 + \frac{\frac{(1-e)\overline{\gamma}}{\overline{H}} - 1}{e\xi_1}\right), 1\right\}, \tag{1}$$

[5] Securitization may also lead to renegotiation frictions for reasons that are orthogonal to the one we propose. For instance, dispersed ownership brought about by securitization or agency problems between services of securitized loans and investors may impede mortgage renegotiation (Piskorski, Seru, and Vig (2010), Agarwal et al. (2015)).

suggesting that the large lender's optimal foreclosure probability λ_1 is less than one if

$$\xi_1 > \frac{\frac{(1-e)\overline{\gamma}}{\overline{H}} - 1}{e}.$$

This condition is more likely to hold, and foreclosures are less likely, if: (1) the share of the market held by the large lender, ξ_1, is sufficiently large; (2) the fraction of households affected by the shock, e, is large; and (3) housing demand relative to the size of the housing stock, $\overline{\gamma}/\overline{H}$, is large.

The existence of a large lender therefore has important implications for the dynamics of the housing market. In particular, it implies fewer foreclosures and higher house prices following negative shocks than in markets with atomistic lenders. In addition, households with high valuations from owning are more likely to remain in possession of their properties even though they suffer a negative income shock.

We now show that this property of the equilibrium generalizes to the general case of N noncolluding "large" lenders.

B.1. The Case of N Large Lenders

Consider N "large" lenders, each with a share, ξ_i, of outstanding mortgages, such that $\xi \equiv \Sigma_{i=1}^{N}\xi_i < 1$. The remaining fraction of mortgages $(1-\xi)$ is held by atomistic lenders, who (as implied by Lemma 1) optimally foreclose with probability one.

Under these assumptions, the market-clearing condition can be written as

$$(1-e)(\overline{\gamma} - \overline{H} - p^{L''}) = \sum_{i=1}^{N} \xi_i \lambda_i e\overline{H} + (1-\xi)e\overline{H},$$

where the right-hand side denotes the supply of foreclosed houses whenever a fraction e of households is hit by a negative income shock, large lenders foreclose with probability λ_i, and atomistic lenders foreclose with probability one. The equilibrium price is now

$$p^{L''} = \overline{\gamma} - \frac{\overline{H}}{1-e} + \frac{\left(\xi - \sum_{i=1}^{N} \xi_i \lambda_i\right) e\overline{H}}{1-e},$$

which is increasing in the overall market concentration ξ and decreasing in the aggregate foreclosure probability, $\sum_{i=1}^{N} \xi_i \lambda_i e\overline{H}$. The optimal foreclosure probability of a lender i is obtained by maximizing the objective function

$$\max_{\lambda_i}(1-e)B + e\lambda_i \left(\overline{\gamma} - \frac{\overline{H}}{1-e} + \frac{\left(\xi - \sum_{i=1}^{N} \xi_i \lambda_i\right) e\overline{H}}{1-e}\right).$$

The first-order condition is

$$\lambda_i = \frac{\frac{(1-e)\overline{y}}{\overline{H}} - 1 + \left(\xi - \sum_{j\neq i=1}^{N-1} \xi_j \lambda_j\right) e}{2e\xi_i},$$

from which the following lemma follows.

LEMMA 2: *The foreclosure probability of lender i, λ_i, is decreasing in the foreclosure probability of other lenders, λ_j.*

The above lemma suggests that each large lender free-rides on the forgiveness decisions of other lenders. Strategic substitutability in foreclosure decisions arises because the equilibrium house price is inversely related to lenders' foreclosure probabilities: if lender $j \neq i$ does not foreclose on a property, the equilibrium house price is, ceteris paribus, higher, making it optimal for lender i to maximize its payoff by foreclosing on delinquent mortgages.

To have a clearer intuition of the equilibrium effects, we consider a symmetric Nash equilibrium with $\xi_i = \frac{\xi}{N}$. Lender i's foreclosure probability is

$$\lambda_i = \lambda = \min\left\{\frac{N}{N+1}\left(1 + \frac{\frac{(1-e)\overline{y}}{\overline{H}} - 1}{e\Sigma_{i=1}^{N}\xi_i}\right), 1\right\}, \tag{2}$$

which continues to be inversely related to the individual share ξ_i of outstanding mortgages. However, relative to the case of one large lender only discussed above (equation (1)), the equilibrium foreclosure probability is higher. The intuition is that increasing the number of large lenders while keeping $\xi \equiv \Sigma_{i=1}^{N}\xi_i$ constant amounts to decreasing the share of outstanding mortgages of each individual large lender, which in turn leads to more foreclosures in equilibrium. Equation (2) also indicates that increasing the concentration of outstanding mortgages ξ, while keeping the number N of lenders constant, decreases the probability of foreclosure.

Importantly, aggregate foreclosures decrease in the concentration of outstanding mortgages. To see this, notice that, with N large lenders, each retaining a fraction ξ_i of outstanding mortgages, aggregate foreclosures are $\lambda\xi e + (1-\xi)e$. It follows that

$$\frac{\partial(\lambda\xi + (1-\xi))}{\partial\xi} = -(1-\lambda) - \frac{N}{N+1}\frac{\frac{(1-e)\overline{y}}{\overline{H}} - 1}{\xi} < 0,$$

suggesting that the aggregate foreclosure rate decreases in ξ.

These results can be summarized in the following proposition:

PROPOSITION 1: *In local markets with N large lenders and a fraction e of distressed households,*

(1) Individual lenders' foreclosure decisions are negatively related to the share of outstanding mortgages that a lender holds on its balance sheet.

(2) *The probability of foreclosure is negatively related to the fraction of local distressed households and positively related to the demand for housing relative to the outstanding housing stock.*
(3) *Individual foreclosure decisions are strategic substitutes. Therefore, each lender's probability of foreclosures increases in the share of the mortgages retained by the other large lenders.*
(4) *In the aggregate, foreclosures are negatively related to the overall concentration of outstanding mortgages on lenders' balance sheets.*
(5) *House prices increase in the overall concentration of outstanding mortgages on lenders' balance sheets.*

In what follows, we bring these predictions to the data.

B.2. Discussion

In the model, lenders with a large share of outstanding mortgages on their balance sheets have an incentive to internalize the effects of foreclosures on house prices. This is the case because their payoffs depend on house prices. The most direct interpretation of this assumption is that the liquidation value of the housing collateral falls as foreclosures increase. However, the externality could operate through different mechanisms. For instance, a generalized decline in house prices may lead to strategic defaults (Elul et al. (2010), Guiso, Sapienza, and Zingales (2013)). Alternatively, a given income shock may be amplified if foreclosures lead to price declines that deteriorate the balance sheets of all households and cause a reduction in real activity, which in turn may lead to further defaults. Also, foreclosures and price declines may impair the balance sheets of lenders if lenders have direct exposures to local real estate through, for instance, the holdings of previously repossessed properties. Externalities may also arise because foreclosures reduce the amenity value of neighboring houses (Fisher, Lambie-Hanson, and Willen (2015)) or because the price declines due to foreclosures prevent borrowers from refinancing. In all of these cases, we would expect lenders with a large share of outstanding mortgages to have stronger incentives to avoid foreclosures.

Our empirical analysis aims to capture any of these mechanisms. Inspired by the equilibrium relations of our stylized model, we estimate reduced-form equations relating foreclosures and house price changes to the share of outstanding mortgages that lenders retain in a neighborhood.

II. Data and Summary Statistics

To test the model's predictions, we use two data sets: a loan-level data set and a zip code–level data set. The loan-level data contain detailed information on lenders, mortgages, and borrowers' characteristics and allow us to study how local variation in the share of mortgages held on a lender's balance sheet affects the likelihood that individual mortgages are foreclosed. The zip code–level data allow us to evaluate the aggregate implications of lenders' incentives to foreclose on local house prices. Below we describe the two data sets.

A. Mortgage Level Data Set

For the purpose of our analysis, a loan-level data set needs to satisfy two requirements. First, it has to contain comprehensive information on the identity of the lender originating a mortgage, the location of the property purchased with a mortgage, and whether a mortgage is securitized or kept on lenders' portfolios. This information is key to measuring the share of mortgages outstanding in a neighborhood that are on a lender's balance sheet and can only be obtained from Home Mortgage Disclosure Act (HMDA) data. HMDA is the largest source of primary U.S. mortgage originations, and covers over 90% of the mortgage activity of commercial banks, thrifts, credit unions, and mortgage companies (see, for example, Mian and Sufi (2009), Avery et al. (2012), Favara and Imbs (2015)).

Second, a loan-level data set must contain information on mortgage performance, including defaults and foreclosures. Unfortunately, this information is not available in HMDA. To overcome this limitation, we follow Agarwal et al. (2015) and merge mortgage originations in HMDA with information on mortgage performance from the Lender Processing Services (LPS) Applied Analytics.[6] LPS contains data on mortgages from the servicers who process mortgage payments and covers about 60% of the mortgage market by value.

Admittedly, the HMDA-LPS merged data cannot be viewed as representative of the U.S. mortgage market, or of portfolio mortgages in particular. The reason is that LPS underrepresents mortgages held by banks on their portfolios, since smaller and midsize banks often service their own mortgages and do not report to LPS. However, since our analysis focuses mostly on variation of portfolio mortgages within lenders and across neighborhoods, there is no reason to expect that this feature of the data should systematically bias our findings.

From HMDA, we select home-purchase mortgages originated in urban areas between 2004 and 2006, and keep track of their performance between 2007 and 2010 using information in LPS.[7] Figure 1 shows the geographical coverage of U.S. counties in the merged HMDA-LPS data set. We are able to merge about 2.5 million distinct mortgages originated by approximately 6,000 lenders in

[6] In performing the HMDA-LPS merge, we replace HMDA lender-identifying information with anonymized identifiers to adhere to the contract terms of the data providers.

[7] All of our results remain unchanged if we consider mortgage performance over the period 2007 to 2009. We include 2010 because foreclosure completions may have been delayed until 2010 in jurisdictions that require the use of a judge to complete a foreclosure procedure (Mian, Sufi, and Trebbi (2015)). The fact that our results are robust to the exclusion of 2010 indicates that the Home Affordable Modification Program (HAMP), introduced at the end of 2009 to help financially struggling homeowners avoid foreclosure by modifying mortgages, is unlikely to influence our results.

Figure 1. County coverage of the loan-level sample.

more than 17,000 zip codes across the United States. Since the main objective of our analysis is to study how lenders behave after borrowers default, we select mortgages that have been delinquent for 90^+ days. We classify a mortgage as delinquent if LPS reports a delinquent status for the mortgage at least once between 2007 and 2010. Similarly, we classify a mortgage as foreclosed if LPS records that a lender has started a foreclosure procedure on the mortgage at least once during the same period. In our sample, about half a million mortgages are 90^+ days delinquent. Nearly 80% of these delinquent mortgages are eventually foreclosed.

We also consider whether a mortgage in default has been modified during the 2007 to 2010 period. We capture mortgage modifications using the algorithm of Adelino, Gerardi, and Willen (2013), which identifies a mortgage modification in LPS whenever the interest rate, maturity, or outstanding balance of a mortgage change, conditional on the mortgage being delinquent. While mortgage modifications may understate lenders' willingness to avoid foreclosures, as lenders may just choose to wait for the borrower to become current again, it is reassuring that the proportion of modifications of 90^+ days delinquent mortgages is slightly above 15%. That is, we can classify nearly the entirety of 90^+ days delinquent mortgages as foreclosed or modified.

We aim to explain differences in the outcome of delinquent mortgages with the lenders' shares of mortgages retained in a neighborhood. As foreclosure externalities may be very localized (Campbell, Giglio, and Pathak (2011)), we use the most granular definition of neighborhood available in HMDA: the census tract. Census tracts are areas designed by the U.S. Census Bureau to be relatively homogeneous in population characteristics, economic status, and living conditions. A zip code typically contains multiple census tracts. We expect foreclosure externalities to arise within census tracts as they have been shown to arise within zip codes (Mian, Sufi, and Trebbi (2015)).

Our main variable of interest is lender l's share of mortgages retained in census tract (i.e., neighborhood) n during the 2004 to 2006 period, which is defined as

$$Retained_{l,n,04-06} \equiv \frac{MR_{l,n}}{TotalOriginations_n}. \tag{3}$$

In equation (3), $MR_{l,n}$ is the number of mortgages retained by lender l in neighborhood n over the 2004 to 2006 period and $TotalOriginations_n$ is the total number of (retained and securitized) mortgages originated by all lenders in neighborhood n over the same period. The numerator includes only mortgages that the lender retains in its portfolio because we wish to measure the credit risk exposure of the lender in neighborhood n.

A mortgage is classified as retained if it is not sold within a year to a third party. Since the process of securitization takes on average two to three months (Rosen (2010)), we consider only mortgages originated in the first three quarters of the year, as mortgages issued at the end of the year may be securitized at the beginning of the following year and thus improperly classified as retained in HMDA.[8] We compute $Retained_{l,n,04-06}$ over the three-year interval from 2004 to 2006 because we wish to measure mortgage holdings in terms of the stock of retained mortgages just before the U.S. foreclosure crisis.

Panel A of Table I provides descriptive statistics for the variables used in the loan-level analysis. The majority of mortgages in the merged HMDA-LPS data set are securitized (about 80%). However, there is also a relatively large sample of portfolio mortgages, which allows us to estimate the effect of $Retained_{l,n,04-06}$ on the probability that a mortgage is foreclosed. As shown in Table I, there is some variation in $Retained_{l,n,04-06}$. While the average $Retained_{l,n,04-06}$ over the full sample including securitized mortgages is small by construction (because securitized mortgages are held by dispersed investors), the average $Retained_{l,n,04-06}$ for portfolio mortgages is 0.03 and is nearly 0.20 in the top percentile.

B. Zip Code–Level Data Set

We also evaluate the aggregate implications of our hypothesis on foreclosure rates and house prices. If foreclosures generate negative externalities that reduce the value of nearby homes, the concentration of outstanding mortgages should correlate not only with foreclosure rates but also with changes in house prices.

To test this hypothesis, we obtain data on house prices from CoreLogic, which provides quality-adjusted price indexes for existing single-family properties at the zip code level. Zip codes are the smallest geographical units with comprehensive house price data. We also obtain data on delinquency rates from

[8] We use HMDA to identify retained mortgages because both the proportion of retained mortgages and the concentration of outstanding mortgages on lenders' balance sheet need to be computed using the universe of mortgage originations.

Table I

Variable Definitions and Summary Statistics

This table provides definitions and summary statistics for the main variables used in the empirical analysis. Panel A provides summary statistics for the 90^{+} days delinquent mortgages in our merged HMDA-LPS data. Panel B provides summary statistics for the zip code–level data.

Panel A: Loan-Level Data Set: Sample of 90^{+} days Delinquent Mortgages in the Merged HMDA-LPS Data

Variable Name	Variable Description	Source	Mean	P10	P90	SD	Obs.
Foreclosure	Dummy variable that takes the value of one if a mortgage is ever foreclosed between 2007 and 2010, and zero otherwise.	LPS	0.747	0	1	0.434	461,149
Modification	Dummy variable that takes the value of one if the mortgage is ever modified between 2007 and 2010, and zero otherwise. Modifications refer to a change in a mortgage's interest rate, principal balance, and loan terms, using the algorithm developed by Adelino, Gerardi, and Willen (2013).	LPS	0.157	0	1	0.364	418,030
Retained	Number of mortgages retained by a lender in a census tract as a fraction of the total number of mortgages originated in the same census tract. Mortgages originated and retained are measured between 2004 and 2006. Lenders include commercial banks, thrifts, credit unions, and mortgage companies.	HMDA	0.00411	0	0.0112	0.0164	461,149
FICO	Borrower's FICO score at origination.	LPS	6.81	6.04	7.58	0.593	461,149

(*Continued*)

Table I—*Continued*

Panel A: Loan-Level Data Set: Sample of 90^{+} days Delinquent Mortgages in the Merged HMDA-LPS Data							
Variable Name	Variable Description	Source	Mean	P10	P90	SD	Obs.
IO	Dummy variable that takes the value of one if a mortgage is interest rate only, and zero otherwise.	LPS	0.302	0	1	0.459	461,149
Mergers	Lender zip code–level measure of mergers between large nonfailing commercial banks (i.e., with assets > \$1bn) during the 2004 to 2006 period, using the share of zip code deposits of the merging institutions as weights.	FRB Chicago & SOD-FDIC	0.00945	0	0	0.0618	461,149
High Cost	Dummy variable that takes the value of one if a mortgage is originated with an interest rate 3 percentage points or more above the rate of a comparable maturity Treasury security, and zero otherwise.	HMDA	0.341	0	1	0.474	461,149
Jumbo	Dummy variable that takes the value of one if a mortgage is a jumbo mortgage, and zero otherwise.	LPS	0.161	0	1	0.368	461,149
LTV	Borrower's loan to value (LTV) ratio at origination.	LPS	0.819	0.747	0.958	0.0918	461,149
Minority	Dummy variable that takes the value of one if the borrower is Black or Hispanic, and zero otherwise.	HMDA	0.384	0	1	0.486	461,149
Securitized	Dummy variable that takes the value of one if a mortgage is securitized within the year of origination through GSE or private label, and zero otherwise.	HMDA	0.844	0	1	0.363	461,149

(*Continued*)

Table I—*Continued*

Panel B: Zip Code–Level Data Set

Variable Name	Variable Description	Source	Mean	P10	P90	SD	Obs.
90^+ Days Delinquency Ratio	Zip code number of mortgages that are 90^+ days delinquent divided by the number of outstanding mortgages. Average from 2007 to 2010.	Equifax	0.0258	0.0078	0.0543	0.0197	6,534
Foreclosure Rate	Zip code number of foreclosures, averaged from 2007 to 2010, divided by the number of single-family, owner-occupied housing units in the zip code in 2000. Foreclosures are defined as the sum of notices of trustee sale and notices of sales.	RealtyTrac.com and U.S. Census Bureau	0.0198	0.0014	0.0489	0.0273	6,159
House Price Change, 2007 to 2010	Logarithmic change between 2007 and 2010 of the zip code house price index for single-family, owner-occupied houses.	CoreLogic	−0.2411	−0.5320	−0.0341	0.1878	6,420
Top 4	Number of mortgages retained by the top 4 holders in a zip code divided by the total number of mortgages originated in that zip code. Mortgages originated and retained are measured from 2004 to 2006. Mortgages are conventional mortgages for the purchase of single-family, owner-occupied houses. Lenders include commercial banks, thrifts, credit unions, and mortgage companies.	HMDA	0.1331	0.0790	0.1997	0.0535	6,570

(*Continued*)

Table I—*Continued*

Panel B: Zip Code–Level Data Set							
Variable Name	Variable Description	Source	Mean	P10	P90	SD	Obs.
Census Income, 2000	Logarithm of median income in the zip code in 2000.	U.S. Census Bureau	10.9509	10.5412	11.3671	0.3279	6,570
House Price Change, 2004 to 2006	Logarithmic change between 2004 and 2006 in the zip code house price index for single-family, owner-occupied houses.	CoreLogic	0.1795	0.0266	0.3568	0.1266	5,547
Judicial Foreclosure	Dummy variable that takes the value of one for zip codes in states with a judicial requirement for foreclosures, and zero otherwise.	Rao and Walch (2009)	0.5339	0	1	0.4989	6,570
Mortgages Originated	Average of the number of mortgages originated between 2004 and 2006 divided by the number of single-family, owner-occupied housing units in the zip code in 2000.	HMDA and U.S. Census Bureau	0.0645	0.0049	0.0471	2.3209	6,185
Population	Logarithm of the zip code's population in 2000.	U.S. Census Bureau	8.4719	8.0953	8.8649	0.3325	6,570
Securitization	Zip code's fraction of mortgages originated for the purchase of single-family, owner-occupied housing units sold within the year of origination to GSE or private label. Average between 2004 and 2006.	HMDA	0.7476	0.6493	0.8260	0.0767	6,570
Zip Code Mergers	Zip code–level measure of mergers between large nonfailing commercial banks (i.e., with assets > $1bn) during the 2004 to 2006 period, using the share of zip code deposits of the merging institutions as weights.	FRB Chicago & SOD-FDIC	0.0296	0	0.0762	0.0402	6,346

Figure 2. County coverage of the zip code–level sample.

Equifax and foreclosure data from RealtyTrac.com. From Equifax, we measure the delinquency rate between 2007 and 2010 as the number of mortgages in a zip code that are 90^{+} days delinquent divided by the number of outstanding mortgages in the same zip code. From RealtyTrac.com, we measure foreclosure rates as the number of properties that receive a notice of sale between 2007 and 2010, divided by the number of single-family, owner-occupied housing units.

Figure 2 shows the geographical coverage of the zip code–level sample. The availability of house price indexes limits the coverage of zip codes relative to the loan-level analysis. However, for this smaller sample of zip codes, we are able to observe almost the entire population of foreclosures.[9] Panel B of Table I presents descriptive statistics for the variables used in the zip code–level analysis.

The aggregate implications of foreclosure externalities depend on the concentration of outstanding mortgages in a neighborhood. Using HMDA data, we construct the index *Top4* as the number of mortgages retained by the four biggest holders in a zip code between 2004 and 2006 divided by the total number of mortgages originated in that zip code during the same period:

$$Top4_{z,04-06} \equiv \frac{MR_{1,z} + MR_{2,z} + MR_{3,z} + MR_{4,z}}{TotalOriginations_z}. \tag{4}$$

In equation (4), $MR_{i,z}$ is the number of mortgages retained by the lender ranked i in zip code z over the 2004 to 2006 period and $TotalOriginations_z$ is the total number of mortgages originated in zip code z by all lenders over the same

[9] RealtyTrac.com is a leading online marketplace for foreclosure properties, covering over 92% of the housing units in the United States.

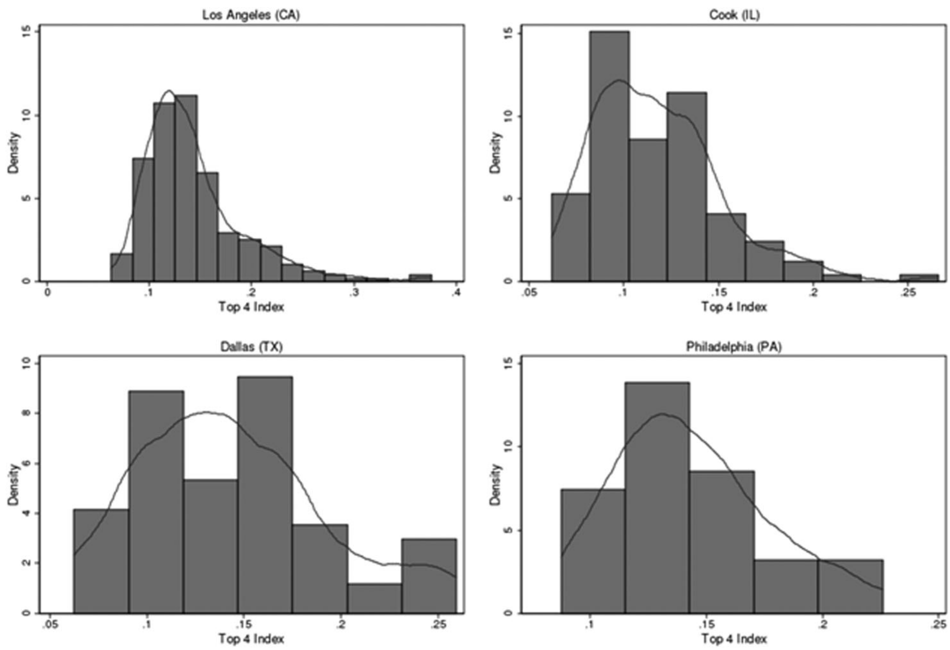

Figure 3. Distribution of the Top 4 index for select U.S. counties (2004 to 2006).

period. As in the loan-level analysis, *Top4* is computed using only mortgages originated in the first three quarters of the year.[10]

As shown in Figure 3, there is substantial nationwide variation in the concentration of outstanding mortgages. In some zip codes, the top 4 holders retain less than 5% of the outstanding mortgages, while in others they retain over 30%. The *Top4* index also shows considerable dispersion *within* counties. Figure 3 displays the distribution of the *Top4* index within three of the largest counties (Los Angeles, Dallas, and Cook) and a relatively smaller one (Philadelphia). As can be seen, the range of the variation in the *Top4* index within these counties is not dissimilar from that observed nationwide.

III. Loan-Level Analysis

A. Empirical Framework

The loan-level data allow us to test whether lenders that retain a high share of outstanding mortgages in a census tract are less likely to foreclose on

[10] A Top 4 index is usually computed to measure concentration of market power (see, for instance, Scharfstein and Sunderam (2015)). Here, Top 4 does not measure market power, but rather the local concentration of holdings of mortgages on lenders' balance sheets.

delinquent mortgages. We estimate the following linear probability model:

$$\Pr(For \mid Delinquency)_{i,l,n,07-10} = \alpha_1 Retained_{l,n,04-06} + \beta X_{i,n,04-06} + \delta_l + \gamma_{n,z,c} + \epsilon_{i,l,n,07-10},$$

where $\Pr(For \mid Delinquency)_{i,l,n,07-10}$ denotes the probability that mortgage i originated by lender l in census tract n is foreclosed during the period 2007 to 2010, conditional on being 90^+ days delinquent. Our hypothesis is that $\alpha_1 < 0$, that is, conditional on default, the probability that a retained mortgage is foreclosed is negatively related to the lender's share of outstanding mortgages in a census tract, $Retained_{l,n,04-06}$.

The regression model is estimated holding constant a vector, $X_{i,n,04-06}$, of mortgage characteristics at origination, such as the borrower's credit score, the loan-to-value ratio, the borrower's ethnicity, the subprime status of the mortgage, as well as its securitization status.[11] The regression model includes also lender fixed effects, δ_l, which ensure that unobserved lender characteristics are unlikely to drive the relationship between $Retained_{l,n,04-06}$ and the probability that a delinquent mortgage is foreclosed. In most of our specifications, we also control for unobserved local market characteristics, $\gamma_{n,z,c}$, using either county (c), zip code (z), or census tract (n) fixed effects.

The linear probability model saturated with fixed effects allows us to evaluate whether the lender's share of outstanding mortgages in a census tract is negatively associated with the probability of foreclosure holding constant unobserved lender and local market characteristics. However, one may wonder whether lenders are more likely to retain mortgages in neighborhoods in which they have an informational advantage over borrowers. This would introduce an endogeneity bias in our analysis. We deem this possibility as unlikely given that our analysis focuses on seriously delinquent mortgages.[12] Notwithstanding, to minimize endogeneity concerns we also estimate the empirical model exploiting arguably exogenous variation in $Retained_{l,n,04-06}$. Following Garmaise and Moskovitz (2006), we exploit mergers of large nonfailing banks of at least 1 billion dollars in assets. The nature and size of these lenders make it implausible that merger decisions depend on housing market conditions in small economic areas, such as zip codes. This ensures that the exclusion restriction is likely to be satisfied. We consider mergers that occurred between 2004 and 2006 and compute the lender's zip code–level exposure to a merger using the share of zip code bank deposits of the merging institutions.[13]

[11] In our analysis, a mortgage is considered subprime if it has high default risk, as measured by an interest rate at origination (from HMDA) that exceeds the prime rate by 3 percentage points or more.

[12] Furthermore, while some lenders may be better at selecting borrowers in a given market, it is unlikely that such an information advantage is stronger in some zip codes or even census tracts. It is therefore reassuring that our results do not change when we include the interaction of county and lender fixed effects.

[13] We obtain mergers from the Mergers and Acquisitions Database of Banks and Bank Holding Companies maintained by the Federal Reserve of Chicago and the zip code address of survivor and

B. Main Findings

Table II reports the main findings. The estimates in column (1) show that, conditional on delinquency, the probability that a portfolio mortgage is foreclosed is negatively correlated with $Retained_{l,n,04-06}$, holding constant borrower and mortgage characteristics as well as lender fixed effects. In addition, column (2) shows that the probability that a securitized mortgage is foreclosed is not related to the share of mortgages retained by the lenders originating a mortgage that is subsequently securitized (*Placebo Retained*). Both results confirm the main predictions of the theoretical model. Since both specifications include lender fixed effects, these results cannot be due to unobservable lender characteristics.

In the rest of the table, we pool portfolio and securitized mortgages to be able to include progressively more granular neighborhood fixed effects.[14] In these specifications, $Retained_{l,n,04-06}$ is equal to zero for securitized mortgages because such mortgages are held by dispersed investors; we also control for a dummy variable that is equal to one if the mortgage is securitized, and zero otherwise.

Across all specifications estimated in Table II, there is a statistically significant negative correlation between $Retained_{l,n,04-06}$ and the probability that the lender forecloses the delinquent mortgage. Importantly, the magnitude of this correlation does not change when we control for unobserved geographical heterogeneity by including county, zip code, or census tract fixed effects in columns (3), (4), and (5), respectively. The stability of the estimated coefficients across specifications suggests that unobservable factors correlated with neighborhood effects are unlikely to be driving our findings.

Table III reports the two-stage least squares (2SLS) estimates. Columns (1) and (2) present the first-stage estimates for the full sample of 90^+ days delinquent mortgages and the subsample of portfolio mortgages, respectively. As expected, $Retained_{l,n,04-06}$ is positively related to the $Mergers_{l,n,04-06}$ instrument, that is, $Retained_{l,n,04-06}$ is higher in zip codes in which lender l ends up with a higher fraction of the deposits as a result of the merger. In both samples, the instruments pass the weak instruments test, as indicated by the Kleibergen-Paap rk Wald F-statistic. The second-stage estimates confirm the negative association between the probability that a delinquent mortgage is foreclosed and the share of mortgages that a lender has retained in a neighborhood.

In columns (3) and (4), the estimates of the coefficient on our variable of interest are larger than the OLS estimates in Table II. This suggests that

nonsurvivor banks' branches, bank assets, and deposits from the FDIC's Summary of Deposits. In our sample, 247 distinct survivoring-large institutions are involved in a merger or acquisition at some point between 2004 and 2006, with branches located in 7,805 zip codes across 940 counties all over the United States. The median (average) large institution has \$365 billion (\$528 billion) in assets.

[14] The inclusion of zip code (census tract) and lender fixed effects amounts to estimating nearly 15,000 (40,000) fixed effects, which implies too demanding a specification given our relatively small sample of portfolio loans. Pooling all loans together (portfolio and securitized), as is common in the literature (e.g., Piskorski, Seru, and Vig (2010)), ameliorates this problem.

Table II

Lender's Share of Outstanding Mortgages and the Propensity to Foreclose

This table provides loan-level OLS estimates of the probability that a delinquent mortgage is foreclosed. The dependent variable is a dummy variable that takes the value of one if a 90^+ days delinquent mortgage is ever foreclosed during the 2007 to 2010 period and zero otherwise. The variable of interest is *Retained*, the share of outstanding mortgages in a census tract retained by a lender during the 2004 to 2006 period. Additional regressors are: dummy variables that take the value of one if a mortgage is securitized (*Securitized*), is a jumbo mortgage (*Jumbo*), is an interest-only mortgage (*IO*), or is a subprime mortgage (*High Cost*), and zero otherwise; the borrowers' loan-to-value ratio (*LTV*) and credit score (*FICO*); and a dummy variable that takes the value of one for Black or Hispanic borrowers (Minority) and zero otherwise. In column (2), *Placebo Retained* is the *Retained* of the lender originating a mortgage that is subsequently securitized. All variables are defined in Table I. Estimates in column (1) are based on the sample of portfolio mortgages only, in column (2) on the sample of securitized mortgages only, and in columns (3) to (5) on the full sample of portfolio and securitized mortgages. Regressions include lender fixed effects and either county, zip code, or census tract fixed effects. Standard errors in parentheses are clustered at the zip code level. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.

	Dependent Variable: Probability that a 90^+ Days Delinquent Mortgage Is Foreclosed				
Sample	(1) Portfolio Mortgages	(2) Securitized Mortgages	(3) All Mortgages	(4) All Mortgages	(5) All Mortgages
Retained	−0.294***		−0.256***	−0.216***	−0.257***
	(0.059)		(0.055)	(0.061)	(0.075)
Placebo Retained		0.021			
		(0.068)			
Securitized			0.030***	0.031***	0.030***
			(0.003)	(0.003)	(0.003)
LTV	0.001	0.073***	0.082***	0.079***	0.079***
	(0.022)	(0.009)	(0.008)	(0.009)	(0.009)
FICO	−0.004	−0.010***	−0.022***	−0.024***	−0.025***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
Jumbo	−0.027***	−0.029***	−0.015***	−0.002	0.001
	(0.005)	(0.003)	(0.002)	(0.002)	(0.003)
IO	0.053***	0.063***	0.056***	0.056***	0.056***
	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)
High Cost	0.060***	0.066***	0.067***	0.063***	0.060***
	(0.006)	(0.002)	(0.002)	(0.002)	(0.002)
Minority	0.001	0.015***	0.004***	−0.000	0.000
	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)
Lender FE	Yes	Yes	Yes	Yes	Yes
Location FE	No	No	County	Zip Code	Census Tract
Obs.	71,977	389,172	461,149	461,149	461,149
R^2	0.084	0.042	0.071	0.101	0.165

Table III

Instrumental Variable Estimates

This table provides loan-level 2SLS estimates of the probability that a delinquent mortgage is foreclosed. Columns (1) and (2) present estimates of the first-stage regression of *Retained*, the share of outstanding mortgages in a census tract that is retained by a lender during the 2004 to 2006 period, on *Mergers*, a bank-zip code measure of mergers between large nonfailing banks during the 2004 to 2006 period. Columns (3) and (4) present estimates of the second-stage of a 2SLS regression in which the dependent variable is a dummy variable that takes the value of one if a 90^{+} days delinquent mortgage is ever foreclosed during the 2007 to 2010 period and zero otherwise. Additional regressors are: dummy variables that take the value of one if a mortgage is securitized (*Securitized*), is a jumbo mortgage (*Jumbo*), is an interest-only mortgage (*IO*), is a subprime mortgage (*High Cost*), and zero otherwise; the borrowers' loan-to-value ratio (*LTV*) and credit score (*FICO*); and a dummy variable that takes the value of one for Black or Hispanic borrowers (*Minority*) and zero otherwise. All variables are defined in Table I. Estimates in columns (1) and (3) are based on the full sample of securitized and portfolio mortgages, and in columns (2) and (4) on the sample of portfolio mortgages only. Regressions include lender fixed effects and county fixed effects. The Kleibergen-Paap rk Wald F-statistic evaluates the null that the instrument (*Mergers*) is excludable from the first-stage regression. Standard errors in parentheses are clustered at the zip code level. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.

	Dependent Variable: Share of Census Tract Mortgages Retained by a Lender		Dependent Variable: Probability that a 90 Plus Days Delinquent Mortgage Is Foreclosed	
	First Stage		Second Stage	
Sample	(1) All Mortgages	(2) Portfolio Mortgages	(3) All Mortgages	(4) Portfolio Mortgages
Retained			−3.647***	−3.334**
			(0.744)	(1.578)
Mergers	0.030***	0.011***		
	(0.005)	(0.003)		
LTV	0.000	−0.010***	0.085***	0.011
	(0.001)	(0.002)	(0.009)	(0.028)
FICO	−0.000***	−0.000	−0.023***	−0.019***
	(0.000)	(0.000)	(0.001)	(0.003)

(*Continued*)

Table III—*Continued*

	Dependent Variable: Share of Census Tract Mortgages Retained by a Lender		Dependent Variable: Probability that a 90 Plus Days Delinquent Mortgage Is Foreclosed	
	First Stage		Second Stage	
Sample	(1) All Mortgages	(2) Portfolio Mortgages	(3) All Mortgages	(4) Portfolio Mortgages
Jumbo	0.000***	0.002***	−0.014***	0.004
	(0.000)	(0.001)	(0.002)	(0.006)
IO	0.000*	0.000	0.056***	0.055***
	(0.000)	(0.000)	(0.002)	(0.005)
High Cost	−0.001***	−0.003***	0.064***	0.050***
	(0.000)	(0.000)	(0.002)	(0.008)
Minority	0.000***	0.001***	0.006***	−0.000
	(0.000)	(0.000)	(0.002)	(0.004)
Securitized	−0.022***		−0.052***	
	(0.000)		(0.018)	
Lender FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Kleibergen-Paap rk Wald F-statistic			32.51	14.32
Obs.	461,149	71,977	461,149	71,977
R^2	0.461	0.349	0.062	0.080

unobserved heterogeneity is likely to bias our OLS estimates downward. But there are also other explanations for the increase in the estimated coefficients. First, our instrument ($Mergers_{l,n,04-06}$) captures only the variation in the share of mortgages retained by the merging institutions. As institutions that merge tend to be more efficient on average (e.g., Erel (2011)), this may help explain why they do a better job of internalizing the adverse spillovers of foreclosures. Second, as the model predicts, lenders' foreclosure decisions are strategic substitutes. Since mergers reduce the number of competitors in a market, the strategic interaction among fewer lenders implies that the size of the instrumental variable estimates should be larger.

That said, the economic effects of the estimated coefficients are plausible. According to the 2SLS estimates, a one-standard-deviation increase in $Retained_{l,n,04-06}$ predicts a decrease in the probability of foreclosure of almost 6 percentage points. This is approximately 10% of the probability of foreclosure for a 90^{+} days delinquent mortgage (which, on average, is 75%). In comparison, in column (3) of Table II, a one-standard-deviation increase in $Retained_{l,n,04-06}$ implies a reduction in the average probability of foreclosure of just 1%.

C. Cross-Sectional Effects

To strengthen the interpretation of our findings, we investigate a few mechanisms leading to the negative association between $Retained_{l,n,04-06}$ and the probability that a delinquent mortgage is foreclosed. We start by testing three comparative statics of our theoretical model, which involve cross-sectional differences between zip codes. Table IV reports the results.

First, our model implies that lenders' foreclosure decisions are strategic substitutes. That is, while lenders with a high share of outstanding mortgages on their balance sheets have weaker incentives to foreclose on defaulting mortgages, these incentives are strengthened when the share of outstanding mortgages retained by other lenders in the same neighborhood increases. The reason is that when other lenders renegotiate defaulting mortgages, housing prices drop less, strengthening a lender's incentives to foreclose so as to maximize the payoff of liquidation. To test this implication, we compute an index of the concentration of mortgages on lenders' balance sheets using a version of the *Top4* index discussed in Section II.B that excludes the share of mortgages retained by a given lender in a zip code if that lender is one of the top 4 holders of outstanding mortgages in that zip code (*Top4 Other*). We then test whether the effect of $Retained_{l,n,04-06}$ on the probability that a delinquent mortgage is foreclosed is weaker in zip codes in which *Top4 Other* is above the county median. The estimates in column (1) of Table IV support the equilibrium implication of our model.

Second, our model implies that the effect of the share of retained mortgages on the probability of foreclosure should be stronger in areas in which the proportion of households that suffers a negative income shock is higher. We test this implication using two alternative proxies for the proportion of distressed

Table IV

Cross-Sectional Differences

This table provides loan-level OLS estimates of the probability that a delinquent mortgage is foreclosed. The dependent variable is a dummy variable that takes the value of one if a 90^{+} days delinquent mortgage is ever foreclosed during the 2007 to 2010 period and zero otherwise. The variable of interest is *Retained*, the share of outstanding mortgages in a census tract retained by a lender during the 2004 to 2006 period. In column (1), *High Top4 Others* is a dummy variable that takes the value of one if the Top 4 index of outstanding mortgage concentration in a zip code (excluding the share of mortgages retained by a Top 4 lender in that zip code) is above the median, and zero otherwise. In column (2), *Prop Affected – Income Shock* is a dummy variable that takes the value of one if a zip code income growth between the 2007 to 2010 and 2004 to 2006 periods is below the median of the distribution in the county, and zero otherwise. In column (3), *Prop Affected LTV* is the zip code share of households with a loan-to-value ratio above 90%. In column (4), *Desirable Neighborhood* is a dummy that takes the value of one if zip code income is in the top quartile of the within-county income distribution, and zero otherwise. Additional regressors are: dummy variables that take the value of one if a mortgage is securitized (*Securitized*), is a jumbo mortgage (*Jumbo*), is an interest-only mortgage (*IO*), or is a subprime mortgage (*High Cost*), and zero otherwise; the borrowers' loan-to-value ratio (*LTV*) and credit score (*FICO*); and a dummy variable that takes the value of one for Black or Hispanic borrowers (*Minority*) and zero otherwise. All variables are defined in Table I. Estimates are based on the full sample of securitized and portfolio mortgages. Regressions include lender and zip code fixed effects. Standard errors in parentheses are clustered at the zip code level. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.

	Dependent Variable: Probability that a 90^{+} Days Delinquent Mortgage Is Foreclosed			
Sample	(1) All Mortgages	(2) All Mortgages	(3) All Mortgages	(4) All Mortgages
Retained	−0.284***	−0.127	−0.090	−0.247***
	(0.069)	(0.078)	(0.093)	(0.058)
Retained × High Top4 Others	0.189**			
	(0.096)			
Retained × Prop Affected – Income Shock		−0.167*		
		(0.093)		
Retained × Prop Affected LTV			−0.493**	
			(0.244)	
Retained × Desirable Neighborhood				0.200*
				(0.114)
Securitized	0.031***	0.031***	0.032***	0.031***
	(0.003)	(0.003)	(0.003)	(0.002)
LTV	0.080***	0.079***	0.079***	0.080***
	(0.009)	(0.009)	(0.009)	(0.008)
FICO	−0.024***	−0.024***	−0.024***	−0.024***
	(0.001)	(0.001)	(0.001)	(0.001)
Jumbo	−0.002	−0.002	−0.001	−0.002
	(0.002)	(0.002)	(0.002)	(0.002)
IO	0.056***	0.056***	0.056***	0.056***
	(0.002)	(0.002)	(0.002)	(0.002)
High Cost	0.063***	0.063***	0.063***	0.063***
	(0.002)	(0.002)	(0.002)	(0.002)

(*Continued*)

Table IV—*Continued*

	Dependent Variable:			
	Probability that a 90^+ Days Delinquent Mortgage Is Foreclosed			
	(1)	(2)	(3)	(4)
Sample	All Mortgages	All Mortgages	All Mortgages	All Mortgages
Minority	−0.000	−0.000	−0.000	−0.000
	(0.002)	(0.002)	(0.002)	(0.002)
Lender FE	Yes	Yes	Yes	Yes
Zip Code FE	Yes	Yes	Yes	Yes
Obs.	461,149	461,149	461,149	461,149
R^2	0.010	0.101	0.0999	0.101

households. In column (2), we focus on zip codes that suffer large negative income shocks defined as zip codes whose income growth between the 2007 to 2010 and 2004 to 2006 periods falls below the median of the county distribution (*Prop Affected – Income Shock*). It appears that the effect of $Retained_{l,n,04-06}$ on the probability of foreclosure is stronger in these more distressed zip codes. Mian, Rao, and Sufi (2013) show that the U.S. housing crisis had bigger consequences in areas with high ex ante leverage. Therefore, as an alternative proxy for the fraction of distressed households, we use the proportion of households in a zip code that during the 2004 to 2006 period borrowed at loan-to-value ratios above 90% (*Prop Affected LTV*). Column (3) shows that the negative correlation between $Retained_{l,n,04-06}$ and the probability of foreclosure is indeed higher in zip codes that are classified as more distressed based on this alternative measure.

Finally, we expect that the effect of $Retained_{l,n,04-06}$ on the probability of foreclosure is lower in areas in which the supply of housing is low relative to demand. To test this prediction, we create a proxy for highly desirable neighborhoods by selecting zip codes that are in the top quartile of the median income distribution in a county (*Desirable Neighborhood*). These are generally zip codes with more amenities, better schools, and higher house prices. If houses are foreclosed in these markets and prices fall, households from other zip codes may move into these highly desirable neighborhoods, mitigating the effect of foreclosures on house prices. Since the externality of foreclosures is expected to be weaker in these neighborhoods, $Retained_{l,n,04-06}$ should have small or no impact on the foreclosure probability. As reported in column (3), we cannot reject the hypothesis that the effect of $Retained_{l,n,04-06}$ is mitigated in highly desirable zip codes.

D. Mechanisms: Mortgage Modifications and Defaults

We also test whether the lower incidence of foreclosures is attained through lenders' willingness to renegotiate defaulting mortgages. In Table V, we

Table V

Mortgage Modifications

This table provides loan-level OLS estimates (columns (1) to (3)) and 2SLS estimates (column (4)) of the probability that a delinquent mortgage is renegotiated. The dependent variable is a dummy variable that takes the value of one if a 90^+ days delinquent mortgage is ever modified during the 2007 to 2010 period, and zero otherwise. The variable of interest is *Retained*, the share of outstanding mortgages in a census tract retained by a lender during the 2004 to 2006 period. Additional regressors are: dummy variables that take the value of one if a mortgage is securitized (*Securitized*), is a jumbo mortgage (*Jumbo*), is an interest-only mortgage (*IO*), or is a subprime mortgage (*High Cost*), and zero otherwise; the borrowers' loan-to-value ratio (*LTV*) and credit score (*FICO*); and a dummy variable that takes the value of one for Black or Hispanic borrowers (*Minority*) and zero otherwise. All variables are defined in Table I. Estimates are based on the full sample of securitized and portfolio mortgages. Regressions include lender fixed effects and either county, zip code, or census tract fixed effects. The Kleibergen-Paap rk Wald *F*-statistic evaluates the null that the instrument (*Mergers*) is excludable from the first-stage regression; the first-stage is reported in column (1) of Table III. Standard errors in parentheses are clustered at the zip code level. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.

	Dependent Variable: Probability that a 90^+ Days Delinquent Mortgage Is Modified			
	(1)	(2)	(3)	(4)
Estimation Method	OLS	OLS	OLS	2SLS
Sample	All Mortgages	All Mortgages	All Mortgages	All Mortgages
Retained	0.170**	0.257***	0.308***	1.384**
	(0.069)	(0.064)	(0.077)	(0.584)
Securitized	−0.014***	−0.012***	−0.012***	0.020
	(0.003)	(0.003)	(0.004)	(0.016)
LTV	0.055***	0.052***	0.045***	0.053***
	(0.008)	(0.008)	(0.008)	(0.008)
FICO	−0.046***	−0.045***	−0.044***	−0.046***
	(0.001)	(0.001)	(0.001)	(0.001)
Jumbo	−0.022***	−0.024***	−0.029***	−0.023***
	(0.002)	(0.002)	(0.002)	(0.002)
IO	−0.022***	−0.021***	−0.021***	−0.022***
	(0.001)	(0.001)	(0.002)	(0.001)
High Cost	−0.032***	−0.030***	−0.028***	−0.032***
	(0.002)	(0.002)	(0.002)	(0.002)
Minority	0.023***	0.022***	0.022***	0.023***
	(0.001)	(0.001)	(0.002)	(0.001)
Lender FE	Yes	Yes	Yes	Yes
Location FE	County	Zip Code	Census Tract	County
Kleibergen-Paap rk Wald F-statistic				30.69
Obs.	418,030	418,030	418,030	418,030
R^2	0.048	0.080	0.149	0.047

estimate the relationship between $Retained_{l,n,04-06}$ and the probability that a defaulting mortgage is modified, using (for brevity) a subset of empirical specifications of Tables II and III.[15] In columns (1) to (3), we estimate the relationship of interest including county, zip code, or census tract fixed effects. In all instances, the correlation is positive and significant as expected.

Column (4) presents the 2SLS estimates. The instrumented effect of $Retained_{l,n,04-06}$ on the probability that a mortgage is modified remains positive and, as was the case for the estimates reported in Table III, the point estimate is larger than its OLS counterpart. In column (4), a one-standard-deviation increase in the instrumented $Retained_{l,n,04-06}$ increases the probability of a mortgage modification by over 2 percentage points—the effect is only half a percentage point in the OLS estimates of column (3).

While in Table V we estimate the probability that a seriously delinquent mortgage is actually modified, lenders that have retained a high share of outstanding mortgages may also be more inclined to renegotiate troubled mortgages before they become seriously delinquent. To explore this possibility, we estimate whether mortgages 30^{+} days delinquent are less likely to become seriously delinquent if they were retained by lenders with high $Retained_{l,n,04-06}$. As shown in column (1) of Table VI, 30-days-delinquent mortgages are less likely to become seriously delinquent when the lender has retained a larger fraction of outstanding mortgages in a neighborhood. Furthermore, column (2) shows that this result is not likely to depend on mortgage quality, as the probability that any mortgage in our HMDA-LPS sample becomes 90^{+} days delinquent does not depend on the share of outstanding mortgages retained by the lender.

However, consistent with our conjecture that lenders with a high concentration of outstanding mortgages are more likely to internalize the spillovers created by foreclosures, column (3) shows that the concentration of outstanding mortgages in a zip code (*Top4*) is negatively correlated with the zip code's delinquency rate. Foreclosures may lead to more defaults not only because a large drop in house prices makes it more difficult to refinance existing mortgages, but also because defaults may affect the social norms for mortgage repayment (Elul et al. (2010), Guiso, Sapienza, and Zingales (2013)). Combined with the result that within a zip code the quality of mortgages held by lenders with a high concentration of outstanding mortgages does not differ from the quality of mortgages held by other lenders, this result fully supports the mechanism behind our hypothesis that a higher concentration of outstanding mortgages leads to the internalization of the externalities caused by foreclosures. It also suggests that these externalities are associated at least in part with the generalized defaults caused by foreclosures.

We next explore an alternative mechanism that may be responsible for the positive correlation between $Retained_{l,n,04-06}$ and foreclosure rates. We ask whether lenders' ability to foreclose may depend on their preexisting organizational capabilities, which may also happen to correlate with their local share

[15] The results are similar if we run regressions on the subsample of portfolio mortgages, though the statistical significance is somewhat weaker.

Table VI

Mechanisms

This table explores the effect of *Retained*, the share of outstanding mortgages in a census tract retained by a lender during the 2004 to 2006 period on mortgage delinquency or foreclosure during the 2007 to 2010 period (columns (1), (2), and (4)) and the effect of *Top4*, the concentration index of outstanding mortgages retained by the top 4 lenders in a zip code during the 2004 to 2006 period on the zip code–level delinquency rate during the 2007 to 2010 period (column 3)). In column (1), the dependent variable is a dummy variable that takes the value of one if a mortgage is ever 90^+ days delinquent during the 2007 to 2010 period and zero if a mortgage is ever 30 days delinquent during the same period; this dummy is not defined for mortgages that never become delinquent. In column (2), the dependent variable is a dummy variable that takes the value of one if a mortgage is 90^+ days delinquent during the 2007 to 2010 period and zero otherwise. In column (3), the unit of observation is the zip code and the dependent variable is the average zip code-level delinquency rate during the 2007 to 2010 period; the regression model in this column includes the following (not reported) zip code–level controls: the fraction of mortgages securitized, the number of mortgages originated divided by the number of single-family housing units, the change in house prices during the 2004 to 2006 period, and the census income and population. In column (4), the dependent variable is a dummy variable that takes the value of one if a delinquent mortgage is ever foreclosed during 2007 to 2010 and zero otherwise. In columns (1), (2), and (4) the variable of interest is *Retained*. In column (3), the variable of interest is *Top4*. In columns (1), (2), and (4), additional regressors are: dummy variables that take the value of one if a mortgage is securitized (*Securitized*), is a jumbo mortgage (*Jumbo*), is an interest-only mortgage (*IO*), or is a subprime mortgage (*High Cost*), and zero otherwise; the borrowers' loan-to-value ratio (*LTV*) and credit score (*FICO*); and a dummy variable that takes the value of one for Black or Hispanic borrowers (*Minority*) and zero otherwise. All variables are defined in Table I. Estimates are based on the full sample of securitized and portfolio mortgages. Regressions include lender fixed effects and location fixed effects as indicated in the table. Each column reports OLS estimates. Standard errors in parentheses are clustered at the zip code level. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.

	Dependent Variable: Probability that a 30-Days-Delinquent Mortgage Becomes 90 Days Delinquent (1)	Dependent Variable: Probability that a Mortgage Is 90 Days Delinquent (2)	Dependent Variable: Zip Code Delinquency Rate (3)	Dependent Variable: Foreclosure and Lender's-County Fixed Effects (4)
Retained	−0.106*	−0.001		−0.239***
	(0.062)	(0.026)		(0.064)
Top 4			−0.084***	
			(0.007)	
Securitized		0.008***		0.026***
		(0.001)		(0.003)
LTV	0.519***	0.217***		0.087***
	(0.016)	(0.003)		(0.009)
FICO	−0.081***	−0.132***		−0.024***
	(0.003)	(0.001)		(0.001)
Jumbo	−0.001	0.016***		−0.017***
	(0.005)	(0.001)		(0.002)

(*Continued*)

Table VI—*Continued*

	Dependent Variable: Probability that a 30-Days-Delinquent Mortgage Becomes 90 Days Delinquent (1)	Dependent Variable: Probability that a Mortgage Is 90 Days Delinquent (2)	Dependent Variable: Zip Code Delinquency Rate (3)	Dependent Variable: Foreclosure and Lender's-County Fixed Effects (4)
IO	0.051***	0.081***		0.054***
	(0.004)	(0.001)		(0.002)
High Cost	0.075***	0.143***		0.066***
	(0.005)	(0.001)		(0.002)
Minority	0.038***	0.067***		0.004**
	(0.004)	(0.001)		(0.002)
Lender FE	Yes	Yes	No	Yes
Location FE	Zip Code	Zip Code	County	County × Lender
Obs.	104,899	2,544,979	6,346	461,149
R^2	0.24	0.256	0.745	0.139

of outstanding mortgages. While it is hard to measure lenders' organizational capabilities, it is reasonable to assume that their infrastructure, technology, and workforce incentives and expectations are not zip code specific, but rather cover larger regional areas, such as counties. To proxy for regional variation in lenders' infrastructure, in column (4), we estimate the probability that a seriously delinquent mortgage is foreclosed, controlling for the interaction of lender and county fixed effects. These fixed effects absorb unobservable lender-county specific factors (including organizational capabilities) that may be correlated with $Retained_{l,n,04-06}$. We find that the coefficient on $Retained_{l,n,04-06}$ is largely unaffected, indicating that differences in lenders' renegotiation/foreclosure capacity are unlikely to be driving our main findings.

IV. Zip Code–Level Analysis

A. Empirical Framework

To evaluate the aggregate implications of our hypothesis, we estimate the following cross-sectional regressions on zip code (z) level data:

$$y_{z,07-10} = \alpha\, Top4_{z,04-06} + \beta\, 90^{+}Delinquencies_{z,07-10} + \gamma\, X_{z,l,04-06} + \delta_C + \epsilon_{z,07-10},$$

where the dependent variable is the foreclosure rate or the logarithmic change in house prices, all measured between 2007 and 2010, and $Top4_{z,04-06}$ is our measure of the local concentration of mortgage holdings between 2004 and 2006. As in the mortgage-level tests, we evaluate outcomes between 2007 and

2010. All controls (with the exception of the delinquency rate) are measured during the period preceding the crisis, that is, 2004 to 2006.

The empirical strategy consists of exploring the effect of *Top4* on foreclosure rates and house prices, controlling for the 90^{+} days delinquency rate, $90^{+}\mathit{Delinquencies}_{z,07-10}$, over the 2007 to 2010 period. The delinquency rate is the single most important determinant of foreclosures and helps us address the main concern that the *Top4* index may be correlated with unobserved factors that affect both foreclosures and house price dynamics. For instance, a negative correlation between the *Top4* index and foreclosure rates could arise because lenders have kept on their balance sheet mortgages originated to more creditworthy borrowers. In this case, ex ante mortgage quality, rather than differences in ex post lenders' incentives to foreclose, could bias our empirical analysis. Similarly, a positive correlation between *Top4* and house prices could reflect heterogeneous income shocks across zip codes over the 2007 to 2010 period that happen to be correlated with our concentration index. The contemporaneous zip code delinquency rate is likely to absorb such confounding effects. While using the 90^{+} days delinquency rate is a coarse control, unobserved heterogeneity across zip codes should not be a big concern because the mortgage-level tests in the previous sections demonstrate that these factors are unlikely to drive our findings.

Our regression framework also includes county fixed effects, δ_C, to control for other unobserved factors that uniformly affect zip codes within the same geographical area, such as state lending or foreclosure laws, or other economic conditions specific to a given county. In addition, since households are likely to approach different lenders within a county to obtain a mortgage, the inclusion of county fixed effects allows us to hold constant ex ante competition in the mortgage market. Finally, the matrix of controls, $X_{z,04-06}$, includes observable zip code characteristics, measured between 2004 and 2006, that are likely to predict our outcome variables. We include characteristics of the mortgage market, such as the proportion of securitized mortgages, the median income, and the proportion of mortgages originated between 2004 and 2006, which may be correlated with the quality of mortgages issued. While these control variables are predetermined, none are truly exogenous. Their inclusion is an attempt to ensure that the *Top4* index has explanatory power, correcting for the usual determinants of foreclosures and house prices.

To provide evidence consistent with a causal mechanism of our main findings, we proceed in two steps. First, we use an instrumentation strategy similar to the one we employ in the loan-level tests. Specifically, we conjecture that the *Top4* index is larger in zip codes that have been more exposed to mergers of nonfailing banks. As in the loan-level analysis, we measure zip code exposure to large lenders' mergers using the share of zip code deposits of the merging institutions. We continue to focus on mergers between nonfailing banks with at least 1 billion dollars in assets because the size and nature of these mergers is less likely to be affected by zip code conditions and therefore more likely to

Table VII
Determinants of the Zip Code Concentration of Outstanding Mortgages

This table provides cross-sectional OLS zip code–level regressions of the *Top4* index of the concentration of mortgages retained between 2004 and 2006 on *Zip Code Mergers*, a zip code–level measure of mergers between large nonfailing banks during the 2004 to 2006 period. All variables are defined in Table I. Regressions include county fixed effects. Standard errors in parentheses are clustered at the county level. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.

	Dependent Variable: Top 4	
	(1)	(2)
Zip Code Mergers	0.116***	0.069***
	(0.030)	(0.017)
90^+ days delinquency ratio		−0.560***
		(0.060)
Securitization, 2004 to 2006		−0.254***
		(0.017)
Mortgages originated, 2004 to 2006		−0.047***
		(0.014)
Census income, 2000		−0.009***
		(0.003)
Population, 2000		−0.024***
		(0.003)
House price change, 2004 to 2006		−0.012
		(0.008)
County FE	Yes	Yes
Obs.	6,346	6,346
R^2	0.622	0.744

satisfy the exclusion restriction.[16] As Table VII shows, our instrument has a strong positive association with the *Top4* index whether we include zip code–level controls or not. Second, we test whether cross-sectional differences in the effect of *Top4* across zip codes with different foreclosure procedures are consistent with a causal effect.

B. Results on Foreclosures

Table VIII reports our main results, with the number of foreclosures per homeowner as the dependent variable. A negative correlation between the *Top4* index and foreclosure rates is consistent with the hypothesis that the local concentration of mortgage holdings mitigates the incentives to foreclose. In column (1), a one-standard-deviation increase in the *Top4* index is associated with a 20% reduction in the standard deviation of foreclosure rates. As in our earlier tests, the effect is larger in column (2), where *Top4* is instrumented with

[16] Our zip code measure of mergers covers a total of 4,984 zip codes across 625 counties in 48 states.

Table VIII
Zip Code Foreclosure Rates and Outstanding Mortgage Concentration

This table provides cross-sectional zip code–level regressions of the foreclosure rate between 2007 and 2010 on the *Top4* index of outstanding mortgage concentration during the 2004 to 2006 period. All variables are defined in Table I. All regressions include county fixed effects. Columns (1) and (3) present OLS estimates. Column (2) reports estimates of a 2SLS regression, for which the first-stage is reported in column (2) of Table VII. Column (4) reports WLS estimates, with observations weighted by the total population. Standard errors in parentheses are clustered at the county level. The Kleibergen-Paap rk Wald F-statistic evaluates the null that the instrument (*Mergers*) is excludable from the first-stage regression. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.

	Dependent variable: Zip Code Foreclosure Rate between 2007 and 2010			
	(1)	(2)	(3)	(4)
Estimation Method	OLS	2SLS	OLS	WLS
Top 4, 2004 to 2006	−0.061***	−0.426***	−0.093***	−0.060***
	(0.013)	(0.144)	(0.020)	(0.013)
Top 4, 2004 to 2006 × Judicial foreclosures			0.052*	
			(0.027)	
90^+ days delinquency ratio	0.772***	0.570***	0.768***	0.776***
	(0.089)	(0.101)	(0.088)	(0.088)
Securitization, 2004 to 2006	0.012**	−0.083**	0.011*	0.012**
	(0.006)	(0.035)	(0.006)	(0.006)
Mortgages originated, 2004 to 2006	0.011	−0.004	0.010	0.011
	(0.012)	(0.012)	(0.012)	(0.012)
Census income, 2000	0.005*	0.001	0.005*	0.005*
	(0.003)	(0.003)	(0.003)	(0.003)
Population, 2000	−0.004**	−0.014***	−0.004**	−0.003**
	(0.002)	(0.005)	(0.002)	(0.002)
House price change, 2004 to 2006	−0.026***	−0.030***	−0.025***	−0.026***
	(0.009)	(0.009)	(0.009)	(0.009)
Kleibergen-Paap rk Wald *F*-statistic		37.917		
County FE	Yes	Yes	Yes	Yes
Obs.	6,016	6,016	6,016	6,016
R^2	0.676	0.072	0.676	0.677

the zip code exposure to mergers of nonfailing institutions: a one-standard-deviation increase in the *Top4* index explains roughly 1.5 standard deviations in foreclosure rates.

Column (3) provides additional evidence consistent with a causal mechanism running from the concentration of outstanding mortgages on lenders' balance sheets to foreclosure rates. In the United States, some states require that the sale of a mortgaged property takes place through court supervision (judicial foreclosure states), while other states give lenders the automatic right to sell the property of the defaulting borrower (power-of-sale states). Foreclosures by power of sale are generally much faster and cheaper than foreclosure by judicial sale (Pence (2006)). Accordingly, lenders' incentives to foreclose should be weaker in judicial foreclosure states regardless of the share of outstanding mortgages on their balance sheets.[17] Consistent with this conjecture, the estimates in column (3) show that the *Top4* index has a muted effect on foreclosure rates in zip codes of judicial foreclosure states. This is an important result as lenders' payoffs in the case of borrower default are likely to be lower in judicial states. Thus, lenders that keep mortgages on their balance sheets should have stronger incentives to originate mortgages of better quality in judicial foreclosure states than in power-of-sale states. This reasoning helps alleviate any residual concerns about endogeneity issues related to the ex ante quality of mortgages.

Finally, in column (4), we report results based on weighted least squares (WLS), with weights given by the zip code population. Heterogeneity in population size may affect the correlation between *Top4* and foreclosure rates because banks may set up renegotiation facilities to deal with a larger number of defaulting borrowers in highly populated zip codes. As shown, we do not find any such effect: the WLS estimates in column (4) are not any different from the OLS estimates in column (1).

C. Results on House Prices

In this final section, we study the implications of lenders' incentives to foreclose on house prices. If an increase in *Top4* is associated with lower foreclosure rates and foreclosures adversely affect local house prices, house prices changes should be positively correlated with the concentration of outstanding mortgages in a neighborhood. Table IX tests this prediction on zip code–level data, with the same empirical framework used in the analysis of foreclosures.

A positive and significant coefficient on *Top4* confirms that, during the 2007 to 2010 period, house prices declined less in zip codes with a higher *Top4* index. In column (1), we estimate that a one-standard-deviation increase in the *Top4* index is associated with 1-percentage-point lower house price depreciation. Once again, the economic effect is larger when we instrument the *Top4* index with the zip code's exposure to large banks' mergers (column (2)). In this case, a one-standard-deviation increase in the *Top4* index is associated with a 6-percentage-point lower house price depreciation, equivalent to one-third of a standard deviation of the change in house prices.

The last two columns report the results of two additional cross-sectional exercises. In column (3), we exploit variation in *Top4* and house prices across states with different foreclosure procedures. As argued in Section IV.B, lenders should have weaker incentives to foreclose if the foreclosure process requires a judicial intervention, regardless of the share of outstanding mortgages on their balance sheets. Our estimates provide support for this prediction: the relationship between *Top4* and house prices changes is weaker in jurisdictions with costly foreclosure procedures. Finally, column (4) presents WLS estimates further showing that our results do not depend on the population of the zip code.

[17] We obtain the list of states in which lenders must receive a judge's approval to foreclose (judicial foreclosure states) from Rao and Walsh (2009).

Table IX

Change in House Prices and the Top 4 Index of Zip Code Outstanding Mortgage Concentration

This table provides cross-sectional zip code–level regressions of the logarithmic change in house prices between 2007 and 2010 on the Top 4 index of outstanding mortgage concentration during the 2004 to 2006 period. All variables are defined in Table I. All regressions include county fixed effects. Columns (1) and (3) report OLS regressions. In column (2), estimates are obtained by 2SLS; the first-stage is reported in column (2) of Table VII. In column (4), estimates are obtained by WLS, with observations weighted by the total population. Standard errors in parentheses are clustered at the county level. The Kleibergen-Paap rk Wald F-statistic evaluates the null that the instrument (*Mergers*) is excludable from the first stage regression. Estimates followed by ***, **, and * are statistically significant at the 1%, 5%, and 10% level, respectively.

	Dependent Variable: House Prices Log Change between 2007 and 2010			
	(1)	(2)	(3)	(4)
Estimation Method	OLS	2SLS	OLS	WLS
Top4, 2004 to 2006	0.119***	0.638*	0.247***	0.125***
	(0.043)	(0.382)	(0.073)	(0.043)
Top 4, 2004 to 2006×Judicial foreclosures			−0.214***	
			(0.082)	
90^+ days delinquency ratio	−1.444***	−1.147***	−1.424***	−1.461***
	(0.165)	(0.240)	(0.161)	(0.162)
Securitization, 2004 to 2006	0.014	0.147	0.016	0.014
	(0.024)	(0.100)	(0.024)	(0.024)
Mortgages originated, 2004 to 2006	−0.010	0.016	−0.009	−0.012
	(0.036)	(0.031)	(0.036)	(0.036)
Census income, 2000	0.014*	0.019***	0.013	0.014*
	(0.008)	(0.007)	(0.008)	(0.008)
Population, 2000	−0.006	0.007	−0.006	−0.004
	(0.005)	(0.010)	(0.005)	(0.004)
House price change, 2004 to 2006	0.033	0.039**	0.032	0.034
	(0.023)	(0.019)	(0.023)	(0.023)
Kleibergen-Paap rk Wald F-statistic		42.398		
County FE	Yes	Yes	Yes	Yes
Obs.	6,346	6,346	6,346	6,346
R^2	0.889	0.083	0.889	0.890

The estimates of *Top4* on foreclosures and house prices allow us to compute an indirect estimate of the zip code–level spillover effects of foreclosures on house prices. We can gauge the zip code–level spillovers of foreclosures to house prices by dividing the coefficient on *Top4* in the house price regressions by the coefficient on *Top4* in the foreclosure regression. The 2SLS (OLS) estimates in column (2) ((1)) of Tables VIII and IX imply that a one-standard-deviation change in the foreclosure rate leads to a decline in house prices of 2.9 (3.9) percentage points. These are conservative estimates of zip code–level externalities of foreclosures. These estimates are smaller but in the same ballpark as those provided by Mian, Sufi, and Trebbi (2015), who document a decrease in house prices between 4 and 11 percentage points for a one-standard-deviation increase in the foreclosure rate.

V. Conclusion

We show that, in markets with a low concentration of outstanding mortgages, lenders exhibit an excessive propensity to foreclose because they do not internalize the effects of foreclosures on house prices. We provide evidence supporting this mechanism during the recent U.S. housing market crisis. We find that not only are lenders less inclined to foreclose in neighborhoods in which they have retained a large share of outstanding mortgages, but also markets with a high concentration of outstanding mortgages on lenders' balance sheets experience fewer foreclosures and smaller house price declines.

These findings have important policy implications. When negative shocks limit borrowers' ability to repay, measures favoring the consolidation of impaired mortgage lenders with similar geographical exposure may increase the concentration of outstanding mortgages. Our findings suggest that these measures may reduce lenders' aggregate losses because they tend to strengthen their incentives to renegotiate defaulting mortgages. Similar effects may be achieved with the creation of bad banks, which, similar to the Home Owners' Loan Corporation during the Great Depression, collect troubled mortgages in times of crisis and are therefore able to internalize externalities in their liquidation decisions.

The mechanism highlighted in this paper has bearings beyond the context of the housing market. In particular, it has implications for the price volatility of any collateralized market with dispersed lending structure. Exploring other areas in which lenders with a high share of the outstanding claims internalize the externalities created by asset liquidations is an exciting avenue for future research.

Initial submission: August 23, 2015; Accepted: August 13, 2016
Editors: Bruno Biais, Michael R. Roberts, and Kenneth J. Singleton

REFERENCES

Adelino, Manuel, Kristopher Gerardi, and Paul S. Willen, 2013, Why don't lenders renegotiate more home mortgages? The effect of securitization, *Journal of Monetary Economics* 60, 835–853.

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, and Douglas D. Evanof, 2015, Loan product steering in mortgage markets, Working paper, University of Chicago.

Anenberg, Elliot, and Edward Kung, 2014, Estimates of the size and source of price declines due to nearby foreclosures, *American Economic Review* 104, 2527–2551.

Asquith, Paul, Robert Gertner, and David S. Scharfstein, 1994, Anatomy of financial distress: An examination of junk bond issuers, *Quarterly Journal of Economics* 109, 625–658.

Avery, Robert B., Neil Bhutta, Kenneth P. Brevoort, and Glenn B. Canner, 2012, The mortgage market in 2011: Highlights from the data reported under the Home Mortgage Disclosure Act, *Federal Reserve Bulletin*, Vol. 98, No. 6, Board of Governors of the Federal Reserve System.

Benmelech, Efraim, and Nittai K. Bergman, 2011, Bankruptcy and the collateral channel, *Journal of Finance* 66, 337–378.

Berger, Allen N., Nathan H. Miller, Mitchell A. Petersen, Raghuram G. Rajan, and Jeremy C. Stein, 2005, Does Function Follow Organizational Form? Evidence From the Lending Practices of Large and Small Banks, *Journal of Financial Economics*, 76, 237–269.

Brunnermeier, Markus K., and Lasse H. Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201–2238.

Campbell, John Y., Stefano Giglio, and Parah Pathak, 2011, Forced sales and house prices, *American Economic Review* 101, 2108–2131.

Elul, Ronel, Nicholas S. Souleles, Souphala Chomsisengphet, Dennis Glennon, and Robert Hunt, 2010, What "triggers" mortgage default, *American Economic Review* 100, 490–494.

Erel, Isil, 2011, The effect of bank mergers on loan prices: Evidence from the United States, *Review of Financial Studies* 24, 1068–1101.

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

Fisher, Lynn M., Lauren Lambie-Hanson, and Paul Willen, 2015. The role of proximity in foreclosure externalities: Evidence from condominiums, *American Economic Journal: Economic Policy* 7, 119–140.

Garmaise, Mark J., and Tobias J. Moskowitz, 2006, Bank mergers and crime: The real and social effects of credit market competition, *Journal of Finance* 61, 495–538.

Gromb, Denis, and Dimitri Vayanos, 2002, Equilibrium and welfare in markets with financially constrained arbitrageurs, *Journal of Financial Economics* 66, 361–407.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2013, The determinants of attitudes towards strategic default on mortgages, *Journal of Finance* 68, 1473–1515.

Guren, Adam, and Tim McQuade, 2013, How do foreclosures exacerbate housing downturns? Working paper, Harvard University.

Harding, John P., Eric Rosenblatt, and Vincent W. Yao, 2009, The contagion effect of foreclosed properties, *Journal of Urban Economics* 66, 164–178.

Hartley, Daniel A., 2014, The effect of foreclosures on nearby housing prices: Supply or disamenity? *Regional Science and Urban Economics* 49, 108–117.

Keeley, Michael C., 1990, Deposit insurance, risk and market power in banking, *American Economic Review* 80, 1183–1200.

Kiyotaki, Nobuhiro, and John Moore, 1997, Credit cycles, *Journal of Political Economy* 105, 211–248.

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

Mian, Atif, Kamalesh Rao, and Amir Sufi, 2013, Household balance sheets, consumption, and the economic slump, *Quarterly Journal of Economics* 128, 1687–1726.

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

Mian, Atif and Amir Sufi, 2011. House prices, home equity based borrowing, and the US household leverage crisis, *The American Economic Review*, 101, 2132–2156.

Mian, Atif, Amir Sufi, and Francesco Trebbi, 2015, Foreclosures, house prices, and the real economy, *Journal of Finance* 70, 2587–2634.

Pence, Karen M., 2006, Foreclosing on opportunity: State laws and mortgage credit, *Review of Economics and Statistics* 88, 177–182.

Piskorski, Tomasz, Amit Seru, and Vikrant Vig, 2010, Securitization and distressed loan renegotiation. Evidence from the subprime mortgage crisis, *Journal of Financial Economics* 97, 369–397.

Rao, John, and Geoff Walsh, 2009, Foreclosing a dream: State laws deprive homeowners of basic protections, Technical Report, National Consumer Law Center, Boston, MA.

Rosen, Richard J., 2010, The impact of the originate-to-distribute model on banks before and during the financial crisis, Working paper, Federal Reserve Bank of Chicago.

Scharfstein, David, and Adi Sunderam, 2015, Market power in mortgage lending and the transmission of monetary policy, Working paper, Harvard Business School.

Shleifer, Andrei, and Robert W. Vishny, 1992, Liquidation values and debt capacity: A market equilibrium approach, *Journal of Finance* 47, 1343–1366.

Shleifer, Andrei, and Robert Vishny, 2011, Fire sales in finance and macroeconomics, *Journal of Economic Perspectives* 25, 29–48.