

Foreclosure Contagion and the Neighborhood Spillover Effects of Mortgage Defaults*

Arpit Gupta[†]

Columbia Business School

January 25, 2016

JOB MARKET PAPER

For the latest version of this paper, visit arpitgupta.info/s/ContagionJMP.pdf

Abstract

I analyze the existence of default spillovers in the residential mortgage market. I focus on shocks to interest rates paid by borrowers resulting from two administrative details in ARM contract terms: the choices of financial index and lookback period. I find that a 1 percentage point increase in interest rates at the time of ARM reset results in a 2.5 percentage rise in the probability of foreclosure in the following year. Instrumenting for mortgage foreclosure using these reset characteristics, I find evidence that each foreclosure filing leads to an additional 0.3 to 0.6 completed foreclosures within a 0.10 mile radius. I document that the price effects of foreclosures on neighboring home prices are unlikely to completely account for the magnitude of this effect. Complementing the price channel, I emphasize two additional mechanisms: a bank-supply channel resulting in a one-third drop in refinancing activity after a foreclosure, and a borrower response channel arising from peer effects. Neighboring borrower payment responses are linked to the timing of local mortgage default, are not associated with defaults on revolving debts, and are concentrated in areas with few nearby local foreclosures—consistent with an information channel based on learning about the costs of default. I also provide suggestive evidence on the macroeconomic impact of foreclosure spillovers: I find that counties and zip codes experiencing greater intensity of reset among borrowers in adverse conditions experience subsequent drops in house prices and higher foreclosure volumes. These results shed light on an important amplification channel of shock transmission during the recent foreclosure crisis.

*I thank Tomasz Piskorski, Edward Morrison, Wei Jiang, Daniel Wolfenzon, Chris Mayer, Emi Nakamura, Patrick Bolton, José Scheinkman, Joe Tracy, Andreas Fuster, Andrew Haughwout, and various colloquium participants at Columbia University and the New York Federal Reserve Bank for helpful comments. I am grateful to Equifax, BlackBox Logic, DataQuick, and Zillow for their data, research support, and infrastructure that were invaluable for the analysis in this paper. The Chazen Institute at Columbia provided critical funding to support this research.

[†]email: agupta16@gsb.columbia.edu. Phone: (828) 280-6638.

1 Introduction

Over 4 million completed foreclosures took place between January 2007 and December 2010, and another 8.2 million foreclosures were initiated in this period (Blomquist 2012). While the foreclosure crisis remains historic in its aggregate cost to homeowners and investors, understanding the precise mechanisms and channels behind this wave of mortgage defaults remains a challenge. In this paper, I examine the role of neighborhood spillover effects from foreclosures as an important amplification mechanism behind this foreclosure crisis.

There are several plausible mechanisms by which foreclosures may affect the default patterns of their geographical peers. Foreclosures reduce the market price of neighboring homes, which in turn may induce those borrowers to default due to negative equity. Alternatively, lenders may react by denying refinancing opportunities to areas which have experienced previous foreclosure activity. A separate channel centers on the role of information: through borrowers' reassessment of the costs of default after exposure to neighboring foreclosures, or through their reassessment of the stigma associated with mortgage nonpayment (as emphasized, for instance, in Guiso, Sapienza, and Zingales [2013]). Finally, a direct treatment effect involves foreclosures inducing greater crime, vandalism, and other forms of property depreciation that reduce the amenity value of the neighborhood.¹ Though the price impacts of foreclosures have been previously studied,² we know comparatively little about the causal implications of foreclosure activity on neighborhood default behavior or the precise mechanisms by which such peer effects operate.

The key contribution of this paper is an empirical setting allowing for the causal estimation of foreclosure spillovers. Understanding this question provides an important context to understanding the seeming snowballing wave of foreclosures observed during the period of initial subprime mortgage defaults and ensuing financial crisis, as well as the policy intervention designed to combat foreclosure externalities. For instance, as Timothy Geithner argued in a speech on February 10, 2009, introducing the Financial Stability Plan, which laid the groundwork for HAMP (Home Affordable Modification Program) and a variety of other government programs: "As house prices fall, demand for housing will increase, and conditions will ultimately find a new balance. But now, we risk an intensifying spiral in which lenders foreclose, pushing house prices lower and reducing the value of household savings, and making it harder for all families to refinance."³ My work examines the existence of foreclosure externalities and the extent to which they constitute an amplification mechanism potentially motivating these federal housing relief efforts.

¹For instance, see Immergluck and Smith (2005, 2006a) on crime and local amenities.

²See, for instance, Campbell, Giglio, and Pathak (2011) or Mian, Sufi, and Trebbi (2015).

³See <http://www.treasury.gov/press-center/press-releases/Pages/tg18.aspx>.

A central econometric challenge to understanding foreclosure contagion is the issue of reflexivity and endogenous assignment of default, as emphasized in Manski (1993). Observing that foreclosures appear to be geographically clustered (for instance, as seems to be the case among foreclosure completions in Phoenix as shown in Figure I) is equally consistent with the possibility that geographically proximate borrowers suffered a common shock (for instance, due to a local plant closing), or share common (possibly unobservable) characteristics that predict mortgage default. Understanding the causal contribution of foreclosures on the default behavior of neighboring properties has proven to be a key challenge in prior literature on this subject.

To address this econometric issue, I provide a novel instrument based on exogenous shocks to interest rates among adjustable-rate mortgage (ARM) borrowers. I argue that these shocks impact the foreclosure of resetting households, but affect the default choices of neighboring properties only *through* the channel of default on treated households. Adjustable-rate mortgages in the U.S.—which were quite common among subprime and jumbo-prime borrowers during the boom in house price appreciation—are characterized by an initial teaser rate that resets to a market interest rate (plus a margin term to correct for risk) after an initial period, typically lasting two, three, or five years.⁴ I focus on two previously unexplored aspects of the mechanics by which ARMs reset: the choices of financial index and lookback date. When ARMs reset, the market interest rate component of the new payment is derived from the prevailing market interest rate according to an index (typically LIBOR or Treasury), taken a certain number of (lookback) days from the reset date. Importantly for my analysis, these interest rates are then fixed for a period of time between six and twelve months after initial reset.

I find variation in these contract terms which drives borrower payment amounts subsequent to mortgage reset. Analyzing the choice of financial index, I find that while LIBOR and Treasury rates tracked each other quite closely prior to the financial crisis, a large spread opened up during the financial crisis—resulting in large differences in payments among borrowers linked to one index instead of another. LIBOR borrowers resetting in January 2009, for instance, paid on average \$11,000 more than otherwise identical Treasury borrowers resetting that same month. I also find that substantial interday volatility in interest rates led to variation in payments paid by borrowers with different lookback terms—15 lookback days instead of 45, for instance. I argue that both forms of interest rate variation are ultimately determined by administrative details of loan contracts, lead to substantial variation in payment terms after reset, and are unlikely to be related to other aspects of loan performance.

In the first stage of my analysis, I find that the size of the within-month interest rate

⁴Initial teaser lengths of one year, seven years, or ten years also exist but are less common.

shock resulting from these contract differences drives default and foreclosure rates among resetting mortgages. The first stage results are substantial and suggest that a 100 basis point rise in interest rates corresponds to roughly a 2.5% rise in the probability of experiencing a foreclosure in the subsequent twelve months, which represents a substantial increase relative to a baseline foreclosure rate of 8%. These results are in line with existing work on mortgage resets (for instance, Fuster and Willen [2015]). I contribute to this literature by obtaining a tighter empirical setting using the within-month variation in interest rates as a shock for the reset window.

My setting provides a clean form of identification for local foreclosure spillovers. I focus on a broad sample of resetting ARM-holders. I develop a novel merge algorithm that links information on these loans—including their contract terms, credit scores, and loan performance—to Deeds records containing the precise geography at the address level in which the borrowers live. I construct neighborhoods consisting of all transacting properties within a radius of 0.10 miles around the resetting ARM-holder to analyze the spillover effects of default.

I contrast my first stage results—which used the interest rate variation induced by the idiosyncratic contract terms such as lookback date and index to predict the default rate among resetting loans—with reduced form estimates, which use the ARM interest rate shocks to predict the default rates among loans *neighboring* those resetting loans. Since neighboring loans are typically set to very different contract terms, there is no ex-ante reason to expect that these exogenous interest rate shocks should predict borrower loan performance in the absence of peer effects or foreclosure externalities. However, my reduced form estimates suggest that higher within-month interest rates among resetting loans also predict foreclosures and defaults among properties in neighboring areas.

Combining the first stage and reduced form estimates, I present an instrumental variables specification in which the interest rate shock of the resetting mortgages serves as an instrument for the foreclosure of the ARM-holder. Estimates using both index and lookback forms of interest rate variation suggest substantial default spillovers. My preferred specifications suggest that each foreclosure raises the probability of default among neighbors by around 2.1%—roughly a 66% impact on the hazard rate of default. Aggregating my estimates at a neighborhood level, these estimates imply that each foreclosure causes an additional 0.3–0.6 foreclosures in neighboring areas. I contrast these estimates with OLS estimates on the impact of foreclosures on neighboring default outcomes, which would suggest much larger estimates around 1—suggesting that each foreclosure is associated with an additional foreclosure in neighboring areas. Finding an IV estimate that is much smaller than the OLS estimate suggests that common local exposure to shocks are an important component to

foreclosure outcomes; but the magnitude and standard errors of my estimate are able to rule out a null hypothesis of zero foreclosure contagion.

My results are subject to a number of important caveats. I assume that within-month variation in interest rates impact neighboring peers of ARM-holders only through the channel of default of the treated properties. This assumption may be violated if, for instance, there is local correlation in contract terms. I test this assumption by focusing on default responses by neighboring fixed-rate mortgage holders, for whom there should in principle be no contagion in default impacts except through exposure to local foreclosures as a result of resetting loan foreclosures.

Importantly, my results are also based on a local average treatment effect (LATE) that is unique to neighborhoods surrounding ARM-holders. I argue this is the relevant neighborhood in several contexts, in particular in evaluating the impact of changes in interest rates on the default behavior of adjustable-rate loans. This population was also the target of many government relief efforts, including HAMP and HARP (Home Affordable Refinance Program), and so understanding default spillovers in this context allows for a better evaluation of federal housing policy targeted toward subprime and other low-income borrowers.

The unusually rich nature of my dataset permits me to go further in analyzing the precise mechanisms responsible for mortgage default contagion and test for confounding factors. I analyze the impact of the price effect, which has been a key focus of prior literature, and find that sales of transacting properties fall by about 1–3% after the foreclosure of a resetting ARM, with smaller estimates when analyzing the repeat-sale estimate of the sale price. These estimates are in line with the prevailing literature on the size of pecuniary externalities resulting from foreclosures. I emphasize that, given prevailing estimates in the literature on the elasticity of default with respect to house prices (e.g., Palmer [2014] suggests that a 10% drop in house prices leads to a 1% increase in the frequency of default), it appears unlikely that the house price channel can fully account for the size of the foreclosure spillovers I observe.

Complementing the pure price channel, I emphasize two other mechanisms of contagion. One is a refinancing channel: I find that refinancing activity drops by one-third among neighboring properties after an instrumented foreclosure. This large change in access to market credit likely reflects in part a bank supply response—banks react to the presence of a nearby foreclosure as an indicator of the likely creditworthiness of local borrowers and cut access to credit.

I also emphasize the role of peer effects and borrower responses on mortgage debt. First, I document that foreclosure spillovers are observed both in defaults (as measured by a notice of default field or payment delinquency) as well as foreclosures. Since foreclosure is a joint

outcome of borrower and lender decisions, this finding suggests that borrower responses play a role in foreclosure spillovers. I also document that mortgage default patterns among neighboring properties appear to decline in the few quarters prior to foreclosure completion—at which point the mortgage default is highly visible in the form of the borrower leaving the house, which is then typically vacant and subject to depreciation and vandalism. Consistent with the idea that active borrower decisions are driving my results, as opposed to common exposure to local shocks, I find that revolving delinquencies do not vary around the window of mortgage delinquency of the resetting loan. I examine the geographical areas most subject to foreclosure spillovers, and find the largest responses in a very tight radius around resetting properties. Finally, I find that foreclosure spillover effects are strongest in areas that have previously experienced few recent foreclosures. These observations are consistent with an information channel of default based on borrowers learning about the costs of default in reaction to a local visible foreclosure.

I conclude by examining the broader macroeconomic impact of resetting ARMs. I create an index of areas ranked by the fraction of mortgages that were (1) resetting under conditions of higher reset interest rates and (2) underwater at the time of reset. Adjustable-rate mortgages facing these conditions default at high rates, as they faced an income shock concurrent with a equity shock that left them without financial incentive to stay in the property nor ability to refinance into a new mortgage (analogous to the classic “double-trigger” conditions of default). I rank geographical areas by the degree to which they contain borrowers facing the joint shocks of negative equity and rate reset. Under conditions of default externalities, the initial wave of defaults resulting from ARMs may spillover to affect other properties, resulting in an amplification of the initial shock to affect a broader set of properties and borrowers. I analyze the subsequent macroeconomic performance of areas prone to double-trigger resets, and find evidence of lower home price recovery and greater foreclosures.

My findings point to the importance of neighborhoods as a source of shock and exposure to information for mortgage holders. I find strong evidence of the role of social influence in default decisions based on local interactions in the context of an identification strategy using within-month variation in interest rates paid by ARM-holders. I find that these rate reset shocks predict default both of resetting mortgage holders as well as those of their neighbors. Overall, these results point to foreclosure contagion representing an important amplification mechanism during the financial crisis in propagating shocks to household defaults.

My paper relates most strongly to existing papers on the local impacts of foreclosures. Many of these examine the price impacts of foreclosures, including Campbell, Giglio, and Pathak (2011), Anenberg and Kung (2014), Immergluck and Smith (2006b), and Gerardi et al. (2012). My work is most closely related to papers examining borrower default spillover

effects from foreclosure, such as Towe and Lawley (2013), Munroe and Wilse-Samson (2013), Agarwal et al. (2015a) and Goodstein et al. (2011). My paper differs by providing a novel instrument to analyze the default decisions of peers and isolating the precise mechanisms behind default contagion.

Additionally, my work relates to a growing literature examining the regional consequences of rate reductions for borrowers and regional outcome variables (due to ARM resets, mortgage renegotiations, and refinancing), including Fuster and Willen (2015), Keys et al. (2014), Di Maggio, Kermani and Ramcharan (2014), and Agarwal et al. (2012, 2015b). My work provides complementary symmetric analysis showing that payment increases (due to rate resets) contributed to the foreclosure crisis and house price declines, and suggests the importance of the contagion channel as an important factor during the financial crisis and recovery.

Other related papers include those on peer effects broadly, including on corporate defaults (Azizpour, Giesecke, and Schwenkler 2015; Benmelech et al. 2014), microfinance debt repayment (Breza 2014) and the broader literature on empirical estimation of peer effects.⁵

The rest of the paper proceeds as follows. Section 2 introduces the data, provides context on ARMs, and introduces the empirical strategy. Section 3 provides main specifications on estimating spillover effects on foreclosures. Section 4 examines some mechanisms for these results. Section 5 examines some robustness checks using the data. Section 6 examines more aggregate responses, and Section 7 concludes.

2 Data and Empirical Strategy

2.1 Data

I assemble a unique and unusually rich dataset for the purpose of this project. I start with an administrative Deeds dataset from DataQuick, which contains geocodes taken at the address level for all transacted properties in select counties. The coverage of this dataset is close to universal among newly originated mortgages for my focus period of 2000–2010, including over 22m transactions in total. I focus on a set of fifty-four counties, which cover roughly 40% of the total mortgage market in the period in question. This dataset includes information at the transactions level on sales, mortgages, and foreclosures on properties. For select counties, the dataset also contains information on notices of default, an indicator of borrowers’ failure to make payments.

⁵See, for instance, Dahl, Loken, and Mogstad (2014), Glaeser, Sacerdote, and Scheinkman (1996), and Bertrand, Luttmer, and Mullainathan (2000)

Using a novel algorithm, I merged this data with BlackBox data on private-label securitized loans, which include loan performance information on default and foreclosure as well as administrative information on loan contract terms. This data has also been merged with credit bureau information from Equifax, including information on credit cards, revolving debt, and other mortgage liens. The appendix describes each dataset and the merging procedure in greater detail. The complete linked dataset allows me to observe mortgage transactions, payment decisions, exact interest rates paid, as well as credit scores and other information on precise geographic addresses.

2.2 Introduction to Adjustable-Rate Mortgages

The typical ARM in the United States is a hybrid ARM. For these loans, the interest rate is fixed for some initial period—often two, three, or five years—and resets on a predetermined schedule for the remainder of the loan. The most common reset frequencies are six months or one year, although other durations exist. At origination, the borrower agrees upon an initial rate and a series of guidelines that determine all future resets. Interest rates upon reset are based on the cost of funds to the bank, proxied by a financial index (typically either Treasury or LIBOR), plus a fixed margin agreed upon at origination. Resets are often subject to other rules, for example, per-period caps on the size of the rate change.

Though 30-year fixed-rate contracts dominate the US mortgage market, ARMs comprise a substantial share of mortgages, and were especially popular in the mid-2000s—especially among subprime borrowers. Figure II illustrates the reset structure of these ARMs. I exploit several features of the contract details of these types of mortgages in the identification strategy.

Resets on ARMs are commonly thought to have been important contributors to borrower default decisions for both resetting loans and neighboring areas, though establishing the relative importance of reset rate shocks remains a challenge. For instance, Smith, Perwien, and Ratcliffe (2009) survey loan counsellors and emphasize the importance of rate reset shocks in explaining delinquency patterns among ARMs. They quote one counsellor as saying “Before the resets they were fine. Once the resets kicked in, then they really started having problems making the payments. Without the resets, we probably would not have the volume that we have today” (p. 7). Sheila Bair, chairman of the Federal Deposit Insurance Corporation noted in Congressional testimony in December 6, 2007, “About 1.7 million hybrid loans worth \$367 billion are scheduled to undergo their first reset during 2008 and 2009. This wave of mortgage resets, in combination with the decline in home prices and limited refinancing options, could prompt hundreds of thousands of additional mort-

gage foreclosures over the next two years. These foreclosures will hurt individual borrowers and their communities, as they potentially could place further downward pressure on home values” (Testimony of Sheila Bair).

A key aspect to the importance of rate reset was borrower unawareness or inattention to specific loan contract terms. For instance, Bucks and Pence (2008) link actual borrower contract terms from lender-reported data with self-reported survey evidence on those terms from the Survey of Consumer Finance. They find that borrowers frequently underestimate and fail to understand the nature of interest rate recast upon reset: for instance, 40 percent of survey respondents report that their interest rates can increase by at most 1 percent or more each period, while fewer than 2 percent of ARM contracts restrict interest rate increases to this level. One rationalization for these findings is a “rational inattention” framework in which the cost of information acquisition regarding precise contract terms is prohibitively expensive relative to the benefits of more accurately forecasting future loan payments. Alternatively, the financial sophistication of borrowers who select or are guided toward ARM products may have been more limited, consistent with a broader literature in household finance that finds borrowers underestimate interest rates and the resulting financial consequences.⁶ Whatever the reason, widespread confusion regarding the precise nature of ARM resets—including the teaser rates, the size of the potential interest rate reset, and prepayment penalties—raises the potential that the timing of ARM reset may represent a particularly important focal period for the timing of mortgage default.

ARM resets have increasingly been the focus of studies of household decision making in the context of interest rate shocks. Fuster and Willen (2015) find downward interest rate shocks lead to a substantial reduction in the rate of mortgage default. Keys et al. (2014) and Di Maggio, Kermani, and Ramcharan (2014) assess how household consumption responds to changes in monthly payments induced by ARM resets. My work builds on this literature using ARM resets as shocks by examining the *within*-month variation in interest rates paid by borrowers upon reset, providing a particularly clean form of identification by which to evaluate the nature of ARM resets. Even if borrowers are not fully attentive to the precise details of ARM reset, some borrowers may respond to the nature of the reset by anticipating its effects and defaulting earlier. Using the surprise or within-month variation in interest rates, as induced by variations in index choice or lookback period, provides a more plausibly exogenous form of variation in interest rate shocks.

It is important to keep in mind that ARM reset shocks may either benefit or shock borrowers depending on prevailing interest rates.⁷ Many ARM borrowers resetting in 2009

⁶See, for instance, Lusardi and Tufano (2008)

⁷For this reason, ARM contracts may be preferred from a monetary policy perspective in enabling maxi-

and 2010, in particular, typically reset to fairly low rates and might well have paid lower monthly mortgage costs relative to fixed-rate borrowers who may have been underwater and therefore unable to refinance to take advantage of lower prevailing interest rates. In addition, teaser rates on many ARM products were fairly substantial as well, and many borrowers of these products were anticipated to refinance into prime mortgages after building a credit history (see Mayer, Pence, and Sherlund [2008] and Foote, Gerardi, and Willen [2012] for a discussion of the role of ARM resets in driving the financial crisis). My identification strategy does not rely on ARM products being unique drivers of the foreclosure crisis as a whole; I focus on ARM resets to illustrate the broader transmission mechanism amplifying the impacts of defaults through spillover effects.

2.3 Identification Strategy

To consider the identification challenge with peer effects in this setting, consider two individuals indexed by i within a variety of groups g , whose default outcomes y_{ig} dependent on each others in the following system of simulations equations:

$$\begin{aligned} y_{1g} &= \alpha_1 + \beta_1 y_{2g} + \delta_1 x_{1g} + \tau_1 x_{2g} + \theta_1 w_g + \gamma z_{ig} + \varepsilon_{1g} \\ y_{2g} &= \alpha_2 + \beta_2 y_{1g} + \delta_2 x_{2g} + \tau_2 x_{1g} + \theta_2 w_g + \varepsilon_{2g} \end{aligned}$$

With additional controls x_{ig} observable characteristics at the individual and group level, and w_g observable characteristics at the group level, and ε_{ig} an error term.

The causal outcome of interest is the average treatment effect:

$$\frac{1}{I} \sum_i \sum_g \beta_i y_{ig}$$

Which is the total number of additional defaults induced by a typical defaulter within her neighborhood.

The key prediction is that foreclosures on properties impose various externalities on their neighbors—through a potential price impact and informational exposure to locally distressed properties—that potentially induce different default behavior among peers. However, the key confounding problem preventing straightforward identification through an OLS regression is the joint determination of defaults through peer effects. Observing a positive correlation between defaults of individuals is consistent with the existence of peer effects by which the default choice of one individual is driven by the default choices of another. Alternatively,

mum pass-through of interest rates to borrowers as in Auclert (2015).

individuals may simply be subject to common group-level shocks—in this context, this may reflect the role of other local shocks such as local plant closings—that affect the default behavior of individuals living in a particular neighborhood. These omitted variables are likely to bias the OLS estimate of the causal parameter of interest— β_i —upwards. This is the standard “reflection” problem as discussed in Manski (1993).

To resolve this identification challenge, an ideal experiment would randomly allocate individuals to neighborhoods of varying foreclosure intensity. I adopt instead a quasi-random approach taking advantage of an instrument that impacts the defaults of particular mortgages, but impacts neighboring mortgages only through the channel of the treated individual. I argue later that the reset structure of ARMs can serve as such an instrument.

2.4 Sample Creation

Figure III illustrates the basics of sample selection. First, I clean and analyze a comprehensive sample of ARMs in my sample of BlackBox private-label securitized loans. For these loans, I observe detailed information regarding the particular contract terms that prevail upon reset, including the date of reset, precise interest rate upon reset, and the choice of financial index and lookback period. I keep ARMs that were current three months prior to their first reset in my main sample.⁸

Next, I link these loans to the broader Deeds sample, which consists of a comprehensive sample of transacting mortgages in a sample of fifty-four counties. Importantly for my purposes, the Deeds sample contains the precise address and geocodes of each property in question. This enables me to construct neighborhoods of 0.10 miles in radius, as illustrated in Figure III, which define the parameters of my final data selection. I use 0.10 miles both to retain comparability with prior literature, as well as because this is a tightly defined geography around the property containing approximately twenty-two local properties.

The empirical strategy requires a computation of distances between households. I use a version of DataQuickDeeds data that include both street addresses and address-level geocodes including the latitude and longitude of each property. I compute the great-circle distances using a standard Haversine formula.

I keep information on all properties in the radius around resetting ARMs, with data taken at the time of reset of the ARM. For my main analysis, I follow Dahl, Loken, and Mogstad (2014) and analyze neighborhoods with one and only one resetting adjustable-rate mortgage, but my results generalize to considering properties neighboring multiple ARMs. For my key dependent variable, I analyze the default and foreclosure behavior of properties

⁸This happens after a period of two, three, or five years.

within the two years following reset of the ARM in the center of each circle. In all analysis, I am careful to exclude the resetting ARM itself from analysis of neighboring outcomes.

I also use this basic sample creation for analyzing other variables and datasets. When analyzing the price impacts of foreclosure, I keep sale data among mortgages that sell at a 0.10 mile radius surrounding the resetting ARM. For a select sample of borrowers with private-label securitized loans, I am able to link up information from BlackBox and Equifax. The resulting sample is essentially a subset of the original BlackBox data that links together loans kept at a 0.10 mile radius. I use this sample to analyze precise information on the payment status of mortgage and revolving debts among neighboring loans.

2.5 Empirical Approach

The empirical strategy follows a two-step regression. In the first stage, I regress the default behavior of a resetting ARM on a variety of controls specific to the loan and geographic area against attributes of mortgage reset. For an ARM i in zip code z and month t , I regress:

$$D_{iz,t \rightarrow t+12} = \alpha + \mu_{zt} + \beta' X_{izt} + \gamma' R_{it} + \varepsilon_{izt}$$

Where $D_{iz,t \rightarrow t+12}$ is the foreclosure on a mortgage i in zip code z in the 12 months after a reset date t .⁹ I include fixed effects for month of reset and zip code X_{izt} , and include a variety of other borrower controls, including: the credit rating of the mortgage (e.g., subprime, jumbo-prime); the level of documentation; the credit score both at origination and reset; the reason for the mortgage (e.g., refinancing or purchase); and the current combined loan to value ratio (including both first and second liens, with a measure of house prices adjusted to the zip code level provided by Zillow).

The instruments are reset shocks controls in R_{it} . In the tightest form of identification, these are the relevant interest rates paid by borrowers of a given index type or lookback period. In the broadest sample, these capture the size of the change in interest rates at the time-of-reset shock. Standard errors are clustered at the census tract-year level.¹⁰

This specification yields the fraction of foreclosures among ARMs that can be attributed purely to the within-month variation in interest rates arising from index choice, lookback

⁹To ensure comparability of results across regions with varying time to foreclosure completion, the definition of foreclosure I use in the first stage is the foreclosure start. Subsequent results focusing on the timing of peer responses suggest that neighboring defaults begin prior to the completion of the foreclosure, suggesting that borrowers potentially respond to the physical abandonment of the property. Selecting foreclosure starts as the dependent variable in the first stage therefore allows for potential peer effects under a broader definition of foreclosure. Results using foreclosure completion as the dependent variable of the first stage yield similar results.

¹⁰Robustness checks with Conley standard errors yield similar estimates for standard errors.

choice, or other forms of reset shocks. As an intermediate step, I also investigate the reduced form specification, which investigates how these reset shocks serve as a treatment for neighboring loans:

$$D_{-i,z,t \rightarrow t+24} = \alpha + \mu_{zt} + \beta' X_{-i,zt} + \gamma' R_{izt} + \varepsilon_{-izt}$$

The reduced form, or intention-to-treat, specification asks how assignment into the relevant treated population—having a neighboring loan that is an ARM with a relatively higher reset shock—impacts the future two-year foreclosure probability.¹¹ The notation X_{-i} indicates that these values are taken from the set of all neighbors of the resetting ARM *except* the resetter itself. While data taken for the resetting ARM are taken from the BlackBox dataset of all private-label securitized loans, information on neighbors is taken from the Deeds dataset, which is comprehensive of all transacting mortgages. Information at the Deeds level on $X_{-i,st}$ includes whether the neighbor itself is an ARM or not; the current combined loan to value of local properties (again including subordinate liens and updating the price through a zip-code level index); and an indicator for whether the property is owned by an investor.¹²

Finally, the instrumental variable specification combines the first stage and reduced form specification into an estimate for how borrowers respond to the foreclosure decisions of their peers as instrumented for by characteristics of the mortgage reset:

$$D_{-i,z,t \rightarrow t+24} = \alpha + \mu_{zt} + \beta' X_{-i,zt} + \delta' \hat{N}_{izt} + \varepsilon_{-i,zt}$$

Where \hat{N}_{izt} captures the share of ARM defaults instrumented for by the within-month difference in interest rates paid by the resetting ARM-holder. The resulting two-stage instrumental variable estimation process calculates the impact of an additional foreclosure, as instrumented for by higher interest rates paid by a resetting ARM-holder, on the foreclosure or default propensity of all local properties. These specifications control for the actual index or lookback period of the resetting loan—allowing for any level differences between loans or neighbors of loans linked to these contract terms. The identification proceeds by using the precise interest rate spread induced by those choices in contract terms to instrument for neighboring foreclosure. The key δ coefficient captures the impact of foreclosure rates as instrumented for using those quasi-exogenous interest rate shocks on neighboring default

¹¹The foreclosure definition used for neighbors in both the reduced form and final IV specifications is the foreclosure completion; which is the only measure available among the full set of neighbors observed in the Deeds dataset. Results using a notice of default field yield similar results

¹²This variable is generated following Chinco and Mayer (2015) by isolating properties in which the mailing address for taxes differs from the property address.

behavior.

This estimation process is subject to a number of assumptions and limitations. First, the exclusion restriction requires that the only pathway by which the instrument can affect properties is through the foreclosure of the treated property. An important test of this assumption will come from local fixed-rate mortgage properties, which pay constant interest rates and so by construction are unlikely to be affected by the interest rate dynamics governing the instrument.

The resulting estimates also reflect a LATE specific to regions surrounding the private-label securitized ARMs analyzed in this project. It is worth bearing in mind, however, that this is the relevant group for a number of interesting analyses. For instance, in analyzing the impact of changes in interest rates on loan performance, ARMs reflect the relevant population exposed to fluctuations in interest rates and their localities reflect the locally treated geographies facing their default choices. Additionally, the sample of subprime and jumbo-prime mortgages in question were highly targeted for a variety of federal programs, such as HAMP and HARP, aimed at mortgage relief at least partially motivated by the issue of preventing default spillovers. Finally, the results, though estimated on a particular sample to achieve greatest identification, illustrate broader mechanisms and processes of social interaction that likely reflect broader patterns in the residential mortgage market.

3 Main Specification—Defaults

3.1 Univariate Comparisons

3.1.1 Index Choice

The first contract detail that I exploit is the financial index that determines the reset rate. There are several potential choices of index, but Treasury and LIBOR rates are dominant, and my analysis focuses on contrasting two of these: the one-year LIBOR rate and the one-year Treasury rate.

Prior to 2007, LIBOR and Treasury rates moved quite closely, with a spread of under 50 basis points. While loans tied to LIBOR typically had slightly lower margins to account for the spread, the difference in risk between the two was perceived to be low and both were viewed as valid proxies for the risk-free rate.¹³ Borrowers may have had some choice of index type; it was typically more often a function of lender intentions on the secondary market.

¹³Hull (2010) notes, "There is a small chance that an AA-rated financial institution will default on a LIBOR loan. However, they are close to risk-free. Derivatives traders regard LIBOR rates as a better indication of the "true" risk-free rate than Treasury rates, because a number of tax and regulatory issues cause Treasury rates to be artificially low."

Many investors in mortgage-backed securities (for instance, European banks) had a cost of funds denominated in LIBOR, and were interested in purchasing assets with a payment structure also determined by LIBOR. In my sample, the loan servicer explains more than 50% of the variance in index; however, many originators provided loans in both categories.

In the pre-crisis period, the precise choice of interest might generate modest differences in interest rates for borrowers in the same product class resetting in the same month, but there was little *ex ante* reason to prefer one over the other. However, as seen in Figure IV, the two rates diverged sharply during the crisis, reaching a maximum spread of over 3 percentage points, and averaging over 1 percentage point between 2007 and 2009. The source of this discrepancy was predominately overnight risk, and this interest rate spread is very similar to the traditional TED Spread (the spread analyzed here differs slightly in that loans were indexed to twelve-month LIBOR and Treasury indices).¹⁴

Summary Table II summarizes information at the level of loan, neighborhood, and neighborhood loans for both LIBOR and Treasury-linked loans. The biggest point of difference is that LIBOR-linked loans tended to have a lower origination balance; Treasury-linked loans were more commonly used above the conforming threshold. Analysis restricting to the jumbo market above the conforming threshold preserves quantitatively similar estimates. Index variation exists only among loans with an initial teaser length of five years; such 5/1 ARMs were disproportionately found among the jumbo-prime market.

First, I illustrate my approach using a simple univariate approach that highlights the index variation in my sample. Recall that ARMs, upon reset, are tied to either a LIBOR or Treasury index. Monthly payments after reset are based on the amortization component of the mortgage, plus an interest rate component that is divided into a margin and index component.

Panel A of Figure V illustrates the consequences of the interest rate divergence between LIBOR and Treasury rates among resetting loans in my sample. The black line illustrates the in-sample interest rate spread among loans in my sample, confirming that LIBOR-linked

¹⁴Another possible source of the discrepancy may have been market manipulation of LIBOR rates. A class action suit filed against the twelve largest banks alleges rate manipulation on LIBOR-linked ARM contracts: “Throughout the Class Period, the LIBOR six-month rates on the first business day of each month are, on average, more than two basis points higher than the average LIBOR six-month rates throughout the Class Period. Additionally, from August 2007 through February 2009, the LIBOR six-month rates on the first business day of each month are, on average, more than seven and one-half basis points higher than the average LIBOR six-month rates. Finally, the LIBOR six-month rates on the first business day of each month are, the great majority of the time, higher than the five-day running average of the LIBOR six-month rate surrounding the first business day submissions throughout the Class Period.” See *Annie Bell Adams, et al. v. Bank of America, et al.* 12 Civ. 7461. However, on balance, banks were likely under-quoting LIBOR rates. Any market manipulation is likely to be uncorrelated with borrower, originator, or neighborhood characteristics around the time of the reset and so is unlikely to impact the identification strategy.

loans resetting in periods of high rate divergence between the LIBOR and Treasury indices did, in fact, pay higher interest rates after the reset. Importantly, these rates were fixed for a period of twelve months subsequent to reset. The component of mortgage payments attributable purely to the index divergence was sizable: in the month of greatest divergence (January 2009), resetting LIBOR-linked borrowers paid over 260 basis points on more an identical contract, originated in the same month, relative to a Treasury-linked borrower. This translates to overall payments of over \$11,000 for the subsequent year of the contract. This represents a substantial payment shock, especially in relation to the sorts of shocks previously studied in the literature.¹⁵

The red line in Panel A of Figure V illustrates the impact of the LIBOR-Treasury spread on subsequent loan performance, graphing the difference between future foreclosure rates on LIBOR-linked loans relative to Treasury-linked loans over the following year subsequent to reset. While the simple univariate graph shows some variation in default patterns unrelated to the underlying index-driven difference, the overall pattern of default divergence follows that of the index and peaks in exactly the month of greatest interest rate divergence (January 2009).

Panel B extends this comparison to examine the behavior of *neighboring* loans residing in neighborhoods around either a LIBOR- or Treasury-linked loan. The nature of sample construction is described above. This graph now illustrates the difference in foreclosure rates among borrowers who live near either a LIBOR- or Treasury-linked loan for the two years following reset of the ARM. The novel finding is that neighbors appear to respond to interest rates among neighboring properties. The month of greatest divergence between loans neighboring LIBOR- and Treasury-linked loans again matches the January 2009 spike in interest rates paid by LIBOR- and Treasury-linked resetting loans.

Overall Figure V provides strong evidence for the role of interest rate shocks at the time of reset as a variable influencing default decisions of borrowers and, consequently, the foreclosure intensity faced by neighbors.

3.1.2 Lookback Period

The second contract detail I focus on is the lookback period. At each adjustment, it is necessary for the lender to compute a new interest rate using the relevant index value. The existence of lookback periods is generally a remnant of pre-computer processing times for updating schedules. To allow for processing and notification of the borrower, lenders typically do not take the contemporaneous value of the index on the day of reset, but instead use the

¹⁵For instance, Parker et al. (2013) analyze the consumption impacts of stimulus checks ranging in size from \$300-\$1,200.

index value at a set period in the past. However, there is no set standard for this period: fifteen, twenty-five, and forty-five days are all common durations. While there is some connection between originator identity and the precise choice of lookback period, there is again no strong *ex ante* reason to prefer one lookback period rather than another and in general the choice of lookback period does not forecast future interest rates.

In periods with relatively low fluctuations in interest rates, the precise choice of lookback period makes little difference. However, when index rates are volatile, differences in lookback period can generate significant differences in the reset rate actually paid *ex post*. Figure VI shows the implications of different lookback periods at a time of significant fluctuation: A loan resetting September 1 2007 could experience a greater than 80 basis point difference in rate depending on the lookback.

Summary Table I summarizes information at the level of loan, neighborhood, and neighboring loans for loans of varying types of lookback periods. Loan variation on lookback period is relatively minor; the point of largest difference is found among loans with a 25 day lookback period, which reset to lower levels of equity than loans of other lookback periods.

3.1.3 Reset Rate Shocks

While the previous two sections investigated particular sources of variation in interest rates paid conditional on month of reset due to particular contract terms, for reasons of power and generalizability in comparisons, I also examine the response of ARM resets in general. Figure VII illustrates a local polynomial smoothing of the effects of all of the within-month variation in reset shocks on default rates. Again, the key comparison is not between loans, which experience an upward reset of, say, 3 percent rather than 1 percent due to resetting in a period of high rather than low interest rates. Instead, the date-demeaned component of reset shocks focuses on the within-month variation in reset shocks arising from a combination of differences in origination rates (between loans with different initial teaser rates), different lookback durations, different teaser lengths, and different index rates. The underlying assumption is that such differences in contract terms driving the subsequent shocks to within-month reset rates offer sufficient variation to examine how borrowers respond to these reset shocks, and so provide an instrument for analyzing neighboring responses to different foreclosure intensity. It is important to keep in mind that whatever factors may have driven borrowers of a particular ARM product to pick a particular contract menu, these decisions are out of the control of neighbors, who must take the contract choices and subsequent default patterns of their neighbors as given.

Panel A of VII illustrates the foreclosure response of an ARM borrower in response to higher reset shocks. Borrowers tend to default more in response to higher-than-average reset

shocks for a given month. Panel B is perhaps more surprising: this graph examines the foreclosure response of neighbors of an ARM product in response to a higher reset shock of the ARM-holder. Neighbors of loans facing higher-than-average reset shock are themselves more likely to experience foreclosure over the subsequent two years following mortgage reset.

3.2 Regression Results

3.2.1 OLS Results

As an illustration, I first examine results that test for geographical clustering in an OLS approach.

In Table III, I present estimates on the relationship between completed foreclosures and the number of additional foreclosure completions around a property through a benchmark OLS calculation. The sample consists of all loans within a 0.10-mile radius around a resetting ARM-holder; and examines foreclosure outcomes neighbors for the two years following the ARM reset. The first column estimates the foreclosure response of borrowers as a function of prior foreclosures in the entire radius prior to reset. Additional controls include a variety of variables taken at the level of both the resetting mortgage and the neighbors. Standard errors are clustered at the tract-year level.

The estimate suggests that a prior foreclosure in the region is associated with a 3% higher probability of subsequent foreclosure among neighboring homes in the vicinity or a 100% increase in the hazard rate of foreclosure. The second column estimates the foreclosure response in the two years after reset among neighbors as a function of all foreclosure of the resetting ARM-holder in the period after reset. The estimate in this column is higher—each foreclosure results in a 4.9% higher chance of foreclosure among neighboring homes. The higher estimate in the second column is likely due to the timing of when foreclosures are estimated to have taken place: the first column examines the post-reset foreclosure response of neighbors in response to a previous foreclosure prior to the reset date; the second column examines foreclosure responses in the post-reset window among both neighbors and resetting loan.

The OLS results, overall, provide suggestive though not causally definitive evidence that defaults may be geographically clustered, and are consistent with a sizable literature examining the existence of geographical clustering of default patterns. The results in column 2 in particular suggest very large apparent contagion effects: given an average of 22 houses in each 0.10 mile radius; those results would suggest that each foreclosure is associated with another foreclosure in subsequent two years.

Figure I provides a visualization of this effect in Phoenix, graphing completed foreclosures

there for 2006–2009. The massive rise in foreclosures, and their spatial clustering in certain areas (for instance, in subprime-heavy Glendale located in the top left part of the city) is certainly consistent with the hypothesis that an element of social interaction regarding mortgage default played a role in the amplification of the wave of foreclosures. However, there are key econometric challenges in a straightforward causal interpretation from the OLS evidence. Local correlations in default rates may simply result from exposure to common local shocks (e.g., a local plant closing) or common unobserved variables (e.g., income).

In other words, the standard endogeneity or reflection problem (as discussed, for instance, in Mankiw [1993]) is particularly problematic in this context. Addressing this issue has proven to be difficult with prior work on peer effects, and in particular on residual mortgage default spillovers.¹⁶ To address this issue, I introduce a novel instrument related to administrative details associated with the reset of ARMs.

3.2.2 First Stage Specification

In Column 1 of Table IV, I estimate the following first stage equation:

$$D_{iz,t \rightarrow t+12} = \alpha + \mu_{zt} + \beta' X_{izt} + \gamma' R_{it} + \varepsilon_{izt}$$

The first row illustrates the impact of a 100 basis point rise in within-month variation in interest rates arising from index choice on the foreclosure probability of the loan over the next year. The results suggest that each 100 basis point rise in interest rates is associated with a 2.5% rise in the probability of foreclosure over the next year, where foreclosure is defined as foreclosure starts. The second row illustrates first stage results using interest rate variation arising from the choice of lookback term, and suggests very similar results. The third row uses all within-month variation in reset rates at the time of reset and yields results that are also very similar.

3.2.3 IV Results

Previous sections discuss the identification strategy. Here, I present evidence on results involving neighbor default as a function of resetter characteristics. Column 2 of IV presents reduced form, or intention-on-to-treat, estimates that examine the effect of higher within-month interest rates conditional on reset on the default behavior of the peers neighbors. The

¹⁶For instance, see Campbell, Giglio, and Pathak (2011), p.15: “Furthermore, foreclosures are endogenous to house prices because homeowners are more likely to default if they have negative equity, which is more likely as house prices fall. Ideally, we would like an instrument that influences foreclosures but that does not influence house prices except through foreclosures; however, we have not been able to find such an instrument.”

reduced form provides a consistent estimate of the impact of having a peer exposed to the treatment group of resetting ARMs paying higher interest rates. Unlike the instrumental variables estimation, the reduced form estimate does not require the monotonicity assumption that higher within-month interest rates conditional on reset did not cause resetters to be less likely to foreclose. This regression includes controls at the level of the resetting loan (including controls for the current combined loan-to-value of the resetting loan, the credit score at origination and just prior to reset, loan attributes such as the purpose and type of loan, as well as an indicator of the servicer); the neighboring loans (including the current CLTV, whether the occupant is an investor, and whether the mortgage is an ARM or a refinanced loan); and the neighborhood (including controls for density, past foreclosures in the area, and the fraction of the population that is nonwhite at the census tract level); as well as zip code and month of reset fixed effects.

Column 2 suggests that higher interest rates paid by resetting mortgages are associated with higher foreclosure rates among geographical peers. Given that these mortgage holders have completely separate mortgage terms from the resetting ARM-holder, and in particular may have a fixed-rate mortgage with constant payments, this apparent correlation points to a strong role for peer interaction effects.

Column 3 combines the first stage and reduced form into an instrumental variable specification in which the foreclosure rate of the resetter is instrumented for by attributes of the reset. The first row, examining peer effects arising from the index variation, suggest estimates comparable to the OLS results, of 0.05. These specifications control for whether or not the neighboring loan is itself LIBOR or Treasury linked; the identification comes the precise interest rate spread induced by the LIBOR–Treasury difference. Under the assumptions of the two stage instrumental variable approach, this would suggest strong foreclosure spillovers leading to a substantial increase in foreclosure intensity as a result of nearby foreclosures.

The second row restricts attention to the variation in within-month interest rates induced by choices in lookback period. These specifications control for lookback dummies as well, allowing for level differences in foreclosure outcomes among neighbors as a result of different lookback choices of the resetting loan. In the first stage, the foreclosure of the resetting mortgage is instrumented for by the variation in interest rates deriving from lookback period. The IV estimate calculates the foreclosure probability of neighbors as a function of this predicted foreclosure rate. The point estimate suggests that foreclosures driven by the lookback interest rate variation result in a 1.6% rise in the foreclosure rate of neighbors in the two years after reset, relative to a somewhat higher baseline rate of 4.1%, corresponding to an aggregate impact of 0.35 additional foreclosures as the result of each resetting foreclosure.

My preferred estimates are in the third row of column 3, which combine all data and use

all within-month variation in interest rates as shocks for the foreclosure of the resetting ARM, and use this instrumented foreclosure intensity to examine the peer effects in foreclosure. The point estimate suggests that foreclosures that are induced by attributes of the reset raise the foreclosure probability of neighbors in a 0.10 mi radius by 2.1%, relative to a mean of 3.4% for the two years following reset. This represents a substantial increase in the hazard rate of foreclosure, though the estimate is substantially smaller than the OLS estimate derived from all foreclosures among the resetting property, pointing to the importance in correcting for the issue of reflexivity in responses to foreclosure. These estimates correspond to a cumulative effect of each foreclosure of 0.46 additional foreclosures in the neighboring area for two years after a completed foreclosure (with a 95% confidence interval of (.25, .67)). These effects are quite large in economic significance and suggest that peer exposure to local foreclosures appear quite large in impacting household foreclosure outcomes.

The last column of IV include the F-statistic for the excluded instruments. Estimates range from 173-1031, which are quite high for this context and alleviate some concerns regarding weak instruments in this context. The strength of this result likely arises from a combination of strong effect sizes in the first stage along with a large sample size.

4 Mechanisms

4.1 Neighboring House Price Effects

In investigating the mechanisms by which foreclosures spillovers may impact neighbor responses, I first focus on the role of house prices. A foreclosed property typically sells quickly through auction at a distressed price. Foreclosures therefore result in an increase in net supply of housing in local areas since defaulters typically rent subsequent to experiencing a foreclosure. Given that residential properties are typically differentiated products, this may result in measurable price impacts in the local neighborhood. Additionally, foreclosed properties typically remain vacant and depreciate through the foreclosure process, resulting in a drop in the capitalized amenity value of local areas.¹⁷ These channels suggest the possibility of foreclosures affecting prices of houses in local areas, which may in turn be driving the default decisions of borrowers. Lower house prices may induce borrowers to default for a variety of reasons, including ruthless default by borrowers in negative equity; some borrowers may be facing other adverse shocks such as unemployment and the additional impact of local house price declines may lead them to default for reasons of a double-trigger default.

¹⁷A price discovery channel may also suggest that the presence of a local foreclosure itself reveals some information about the future price dynamics of a neighborhood.

Figure VIII uses an event study framework to illustrate the impact of a foreclosure on sale prices of houses within a 0.10 mile radius around the foreclosed property both before and after the foreclosure completion. The estimates are drawn from a regression of the log sale price of properties against a variety of household characteristics:

$$\log(P_{it}) = \alpha + \mu_t + X_{it} + \sum_{k=-s}^{s-1} \mu_k \cdot 1[(t - T_i) = k] + \varepsilon_{it}$$

Where $\log(P_{it})$ is the log of the sale price of homes, fixed effects control for the month of observation, the zip code, and a variety of property-specific controls. The key coefficients of interest are μ_k , which measure the change in sale price of the house during s quarters prior to and following the timing of the foreclosure in quarter T_i . This sample restricts to borrowers in the vicinity of a resetting ARM that experiences foreclosure. Standard errors are clustered at the tract-year level. Panel A illustrates the impact of the foreclosure on the log price of transacted properties. The estimates suggest a sizable drop in the sale price of transacted properties around 1–4% after the sale. These estimates are comparable to those found in prior research, such as Anenberg and Kung (2014) or Campbell, Giglio, and Pathak (2011), which estimates a roughly 1% impact on price after also differencing out price impacts on more distant geographies. Panel B estimates the impact of foreclosures on the repeat-sale measure of prices, in which the dependent variable is now $\log(P_{i,t-q_i})/\log(P_{it})$, where $P_{i,t-q_i}$ is the last transacted sale on the property. The repeat-sale measure of price, as emphasized in Fisher, Lambie-Hanson, and Willen (2013), allows for a measure of price impact that better controls for property-specific characteristics. The impact of foreclosures on the repeat-sale measure of price is lower—averaging around 1% and reaching statistical significance and a minimum of around 2% two quarters after the completed foreclosure sale.

The difference between the impact of the foreclosure on the standard price relative to the repeat-sale measure of price suggests an important role for compositional effects subsequent to foreclosure completions on the set of properties that choose to transact. Lower-priced properties disproportionately transact subsequent to foreclosures, suggesting a larger price impact than if changes in household sale choices were held fixed. Consistent with this interpretation, figure VIII also illustrates in the background a histogram of completed sales in the neighborhood around foreclosed properties. Properties disproportionately do not transact in the quarter after a completed foreclosure has taken place, in the time when price effects are also greatest.

Addressing the impact of changes in composition on the causal estimation of prices is a challenge that prior work in the area has not fully addressed. However, I emphasize several features of this bias. It appears that much of this selection effect appears to come from

higher priced properties choosing *not* to transact after a nearby foreclosure, suggesting a downward bias on estimated price effects. Second, precise identification of the true treatment of foreclosure effects is not necessarily critical in the context of establishing peer effects, because nearby borrowers and lenders themselves take house prices as given and do not necessarily filter the true price impact separate from the change in composition, which is a challenge even for the econometrician. In particular, property value assessments take local comparable sales as given when establishing estimates of current market prices.

Finally, I emphasize that price effects appear to be unlikely to explain the full magnitude of the foreclosure spillovers I observe, suggesting a role for peer effects operating in conjunction with the price channel. Assuming a price effect of 1%—consistent with the prior literature and my repeat sales evidence—along with conventional estimates of the elasticity of default with respect to price (Palmer [2014] suggests that each 10% drop in house prices results in a yearly increase of mortgage default of 1%) would suggest that the price channel can explain perhaps 10% of the foreclosure contagion channel I observe. Given the difficulties in accurately measuring the price channels, I emphasize the order of magnitude of this result: even under reasonable adjustments to the precise estimation of the price effect, it seems unlikely that the price effect can explain the entirety of the foreclosure contagion channel.

4.2 Refinancing Channel

Supposing that the price channel—a key focus of prior research in the area of foreclosure impacts—is unlikely to explain the full magnitude of my foreclosure contagion results, what are other likely channels of impact between foreclosures and neighbor responses? One important channel complementing the role of house prices is the shift in access to new credit as suggested by changes in refinancing activity.

Table V estimates a two-stage instrumental variable regression in which the outcome in the second stage is refinancing behavior among neighboring loans (again within a 0.10 mi radius); foreclosure of the resetter remains the dependent variable in the first stage. While refinancing opportunities respond in part to changes in house prices as captured in the previous section, separate estimation of the refinancing channel suggests a more precise isolation of this exact pathway. In particular, lenders may respond to both the price effect as revealed through changes in transacting prices, as well as to actual evidence of foreclosure activity in nearby areas. This may be because lenders view foreclosures by themselves as revealing information regarding the quality of the neighborhood or the borrowers contained within it.

I find substantial evidence that refinancing activity drops in areas that have experienced foreclosures, using the instrumented value of foreclosure among resetting loans from the variation in reset rates. Column 1 estimates refinancing results on the BlackBox dataset, for which refinancing is defined using an early prepayment indicator drawn from a sample of private-label securitized loans.¹⁸ Additional controls in BlackBox specifications include a full set of controls for loan characteristics that are commonly measured on both resetting and neighboring loans, such as interest rates, credit score, loan purpose, and equity. My estimates on this sample suggest that instrumented foreclosure activity lowers the frequency of prepayment by 3.5%, or lowering the the hazard rate of prepayment among securitized loans by almost a half.

Effect sizes are larger among the full sample of Deeds loans, shown in column 2, for which an indication for refinancing is drawn from mortgage transactions in the Deeds records. Depending on the specification (relying on index, lookback, and all interest rate variation), I estimate between a 6–8% drop in the future two-year propensity to experience a refinancing, relative to mean of around 17%, or roughly a third drop in the hazard rate of refinancing.

This refinancing effect may reflect either a borrower-driven response if the local foreclosure reduces borrower propensity to seek out new refinancing, or a lender-driven response to the extent that lenders perceive local price drops or other foreclosure externalities that make them more reluctant to extend credit in the vicinity of a foreclosed property. However, I emphasize that the coefficient size of this effect is much larger than my estimates for foreclosure spillovers. That suggests that even if all borrowers who ceased to make payments and wound up in foreclosure would have counterfactually received a new loan instead; there remains a substantial group of borrowers who neither refinance nor default. Instead, they simply continue to make payments on their old mortgage.

The presence of these borrowers is strongly indicative, though not fully conclusive, of a lender-driven response. These borrowers are plausibly denied credit due to the local foreclosure. While some borrowers denied refinancing default as a result, others continue mortgage repayment. A lender-driven response is consistent with many borrowers being locked out of credit markets but continuing to make payments; it is more difficult to rationalize a demand-driven response by which the exposure to a nearby foreclosure would induce borrowers to simply refuse available new mortgage credit which in this period typically lowers their monthly mortgage payments by a substantial amount.

A substantial lender-driven refinancing response helps to reconcile the magnitude of my

¹⁸A downside of this measure is that it also includes early payment of the loan associated with the borrower prepaying and moving. Additionally, the sample of private-label securitized loans during this period typically faced difficulties in accessing refinancing markets throughout this period.

foreclosure spillover effects. If lenders react to local foreclosure activity heavily by denying credit in substantial amounts, neighboring borrowers will find it relatively much more difficult to make mortgage payments and will experience higher rates of default.

4.3 Neighboring Peer Responses

4.4 Estimating Default Outcomes

A separate channel which can also explain the magnitude of my results is a neighboring peer effect channel in which borrowers react to local foreclosure by adjusting their own payment behavior. I provide a range of results illustrating the impacts of local foreclosures on borrower behavior.

Table VI focuses on outcomes of default among neighbors in the second stage, while still using the foreclosure of the resetter in the first stage. Since foreclosure is a joint decision between borrowers and lenders, one potential concern might be that apparent spillover effects on foreclosures may be a lender-driven response. Column 1 assesses this channel by using the notice of default as the dependent variable. This is a relatively automatic public notice delivered to borrowers that they are behind on payments, and have some period of time to resume payment or lenders may initiate a foreclosure process. I find sizable impacts of foreclosures of the resetting mortgage, as instrumented for by reset characteristics, on the default decisions of neighbors. However, since I do not have notice of default information for as many counties, the sample size is somewhat smaller and the lookback and index specifications do not reach statistical significance.

Column 2 of this table examines a variety of outcomes for the BlackBox sample. This sample consists of loans in the BlackBox dataset that match to properties within 0.10 miles of a resetting ARM. While this sample is restricted in terms of coverage of which loans are present, I am able to control in this group for a variety of additional characteristics of loans. In particular, I control for the same characteristics as for the resetting loan, including: the precise contract terms relating to reset (including whether or not the loan is an ARM, the exact interest rate, time from reset, as well as the index type and assigned lookback), along with a variety of other loan-level details such as the combined current loan to value (including both first and second liens, with the home value adjusted to the monthly level using Zillow zip-code data), and the purpose and type of the loan. I also include additional borrower-level data such as the lagged credit score and credit utilization. A key advantage of the BlackBox specification is that I am able to fully control for the same set of covariates for both resetter and local borrower, allowing me to control for any local correlation in loopback periods or index choice.

My estimates suggest that completed foreclosures have a substantial impact on foreclosure start (a coefficient of 0.064) and delinquency outcomes (a coefficient of 0.076) among private-label securitized loans. These effects are sizable and isolate the borrower delinquency, as opposed to purely the lender decision to foreclosure, as a key driver of foreclosure spillover effects. While the implied aggregated default estimates in this sample are quite large (i.e., the total spillover effects on foreclosure start on the BlackBox sample would suggest an additional 1.4 foreclosure starts total in neighboring areas); these estimates entail a hypothetical neighborhood consisting of only subprime or jumbo-prime loans. Neighborhoods typically contain a mix of subprime and other borrowers; an advantage of calculations on the Deeds sample is that this sample consists of the universe of all neighboring borrowers, enabling greater generalizability.

4.5 Timing of Default Responses

I examine how repayment of mortgage debt varies around the timing of the local foreclosure using an event study framework. This connects the precise payment decisions of borrowers to the foreclosure decisions of their neighbors. To evaluate the default decisions of neighbors, I examine the payment behavior of neighbors who hold a mortgage in the BlackBox sample, restricting attention to those borrowers with privately securitized mortgages. I run the regression:

$$O_{it} = \alpha + \mu_t + X_{it} + \sum_{k=-s}^{s-1} \mu_k \cdot 1[(t - T_i) = k] + \varepsilon_{it}$$

Where O_{it} is the outcome of making payments on either a mortgage or other revolving debt (e.g., credit cards), fixed effects control for the month of observation, the zip code, and a variety of property-specific controls. The key coefficients of interests are μ_k , which measure the change in sale price of the house during s quarters prior to and following the timing of the foreclosure in quarter T_i . This sample is restricted to borrowers in the vicinity of a resetting ARM that experiences foreclosure. Standard errors are clustered at the tract-year level.

Panel A of Figure XI illustrates how borrower default patterns behave around the timing of the foreclosure completion of the resetting ARM by plotting the coefficients μ_k . I find that borrowers are less likely to be current the quarter prior to foreclosure completion, during which time the property is typically vacant and the nature of the foreclosure process is visible to neighbors. Mortgage repayment rates continue to decline in the quarter in which the foreclosure is completed and flatten out thereafter.

By contrast, Panel B of Figure XI illustrates borrower repayment on *non*-mortgage debt

as captured by payment status on revolving debt, which includes credit cards. In contrast to Panel A, Panel B shows no impact of the neighboring foreclosure on the payment status of non-mortgage debts. Both the timing of the mortgage default decision—around the moment in which the foreclosure becomes visible to neighbors—as well as the fact that it leads borrowers to default on mortgage but not non-mortgage debts—is suggestive that borrower responses specifically tied to mortgage non-repayment dominate the spillover effects of foreclosures. This result suggests that borrower attitudes specifically towards mortgage debts, as opposed to loans in general, are driving my results.

4.6 The Pattern of Peer Strategic Responses

Table VII cuts my main specification along several dimensions. I follow the instrumental variable approach as outlined in column 3 of table IV, using all of the within-month variation in interest rates and focusing on a variety of samples. First, I find that my results persist when subsetting on neighbors with fixed-rate mortgages. This specification avoids potential pitfalls relating to correlation in loan contracting terms across neighbors, as it is difficult to think of why contract decisions resulting in different interest rates for adjustable-rate borrowers should have any impact on neighbors with fixed-rate contracts. The effects remain sizable among borrowers with fixed-rate contracts—I find an increase in the future two-year foreclosure probability of 1.9%, corresponding to a cumulative impact of an additional 0.42 foreclosures in a hypothetical neighborhood consisting of only fixed-rate mortgages.

I also find larger effects among units in apartments, and sizable effects on investor properties¹⁹. I find very large effects of a 3% rise in foreclosure among underwater properties, though this specification is measured only on the BlackBox sample and does not reach statistical significance.

While failing to repay bills seems an intuitive option, the household financial literature has documented a relative unwillingness on the part of households to exercise their default option when doing so would improve household prospects. For instance, Fay, Hurst, and White (2002) find that a substantial fraction of households for whom it would be financially profitable to declare bankruptcy choose not to do so. Among mortgage holders, the majority of those facing negative equity—owing more on their mortgage than it is worth—continue to make mortgage payments, even though it would be financially profitable for them to cease payments and walk away from the property.

To be sure, the decision to continue making payments on an underwater property is a complicated one involving a variety of other factors. Borrower repayment may be rationalized

¹⁹Where investor status is identified by contrasting the address field in Deeds records across the mailing address where tax bills are sent and the property address as in Chinco and Mayer (2015)

under a dynamic model incorporating the future credit costs of default, the local amenity values of the neighborhood, transaction costs in finding a new residence, or a variety of other factors (though the prospect of being able to live rent-free in the property subsequent to default is a factor which should encourage an earlier default). Nevertheless, it seems likely that many borrowers are making decisions to continue mortgage payments that may be narrowly unprofitable from the perspective of maximizing the default option on the property.

In that context, my preferred interpretation is that default spillovers are high among borrowers for whom it is profitable to default, and who have been previously dissuaded from defaulting for reasons of social stigma or misperception regarding the viability of the default option. The physical presence of a local foreclosure, in this context, appears to trigger additional defaults among these borrowers.

Figure IX illustrates default spillover effects by the year of observation. I focus on the year of reset of the ARM, and follow the subsequent two-year foreclosure probability of neighboring loans using the baseline instrumental variable specification focusing on all loans and variation induced by the reset amount. I find that spillover effects seem fairly stable over the period 2006–2010, and cannot reject a null that they are all the same.

Figure X examines results by distance from the resetting property. I focus on a sample of BlackBox loans within 0.15 miles of a resetting loan, and compute the baseline instrumental variable specification as illustrated in column 3 of Table IV for properties at various distances away from the resetting loan. The estimated relationship shows a roughly linear relationship between distance from the resetting loan and the size of the foreclosure spillover effect. Among properties 0.10–0.15 miles away from the resetting loan, the size of the spillover effect is no longer statistically significant.

Table VIII shows spillover effects ordered by the intensity of prior foreclosures in the area. I find that foreclosure spillovers are strongest in areas that have only experienced zero or one foreclosure within a broader radius (0.25 miles) in the previous two years. This pattern of results suggests that foreclosure spillovers are driven by learning about the cost of default, which manifests most strongly upon the initial set of foreclosures. An alternate hypothesis would be that foreclosures impact peer responses through a channel of affecting the stigma attached to repaying debts, which might suggest that an increasing wave of local foreclosures is required to induce other borrowers to repay. I emphasize that the standard errors are large in the context and I am not able to rule out a more S-shaped pattern of foreclosure spillovers. However, the point estimates I find are most consistent with a peer effect channel which strongly weighs the impact of initial local foreclosures.

The overall pattern suggested by these separate cuts is suggestive of a neighboring peer default channel. I document that the default responses to a local foreclosure (1) is linked to

borrower decisions to cease debt payments, (2) is linked to the timing of serious delinquency on neighboring properties, (3) does not entail delinquency on non-mortgage debts, (4) is highest among underwater properties, (5) is closely tied to the geography of foreclosed properties at a very tight radius, (6) and is highest in areas which have previously experienced little foreclosure activity. The balance of all of these factors suggests a responsive default motive on the part of local borrowers, who learn from local foreclosure activity about the cost of default and cease debt repayment on their mortgage when it is profitable to do so.

5 Robustness

5.1 Placebo Test: Prior Foreclosures

A key identifying assumption is that within-month variation in interest rates, conditional on other observables (such as zip code and characteristics of both the resetting and neighboring mortgages) have no impact on neighboring mortgage default patterns *except through* the channel of foreclosures on the resetting ARM. While this assumption cannot be tested directly, I provide some evidence in its support. One such test is in Table X, which finds that default spillover effects are concentrated among loans within a very small radius directly surrounding the resetting loans. While it is likely that mortgage holders are commonly exposed to local shocks—for instance, the closure of a local plant—it is more difficult to imagine real economic shocks that are correlated with within-month variation in interest rates as induced by the lookback period or index type, but only at extremely small distances around resetting ARMs.

Another robustness check is in Table X, which performs a similar regression as in the main specification—but with the key dependent variable being foreclosure status among loans neighboring the resetting ARM in the *prior* two years before the resetting ARM. This serves as a placebo check of the main analysis: if the instrument is valid, it should not predict foreclosures prior to the time when the resetting ARM is actually paying the differential interest rates. Table X provides evidence consistent with this intuition. Column 1 shows estimates for the BlackBox sample of loans neighboring the resetting ARM; Column 2 provides evidence on the full Deeds sample. Estimates of the instrumented peer effect are insignificant for the BlackBox sample, and significant but negative and very small for the Deeds sample (-0.0005). If anything, the estimates from the prior timing regressions suggest that the effect size may be slightly underestimated due to the assignment of treatment in areas with a slightly lower propensity for foreclosure.

6 Aggregate Effects Reset Intensity and Foreclosures

The previous sections focus on micro-level data to evaluate the effects of foreclosure spillovers on local outcomes. However, many of the effects of resets and foreclosures can be expected to aggregate and impact broader areas. While identification is more difficult when examining these aggregate effects, a key motivation in understanding foreclosure spillovers is in evaluating the broader macroeconomic consequences of household default decisions. The aggregation of externalities resulting from foreclosure presents a possible motivation for a variety of federal debt relief programs—including federal assistance in mortgage modification, refinancing, as well as the central bank’s choice of interest rates, which directly feeds into the payment decisions of ARM-holders.

I analyze the aggregate effects of foreclosures by focusing on regional variation in the fraction of local resetting ARMs. Since interest rates paid at the time of reset can be either higher or lower than the original rate, and even interest rate shocks may not be binding for borrowers with equity, I focus attention on what I refer to as double-trigger ARMs: those facing both (1) an interest rate hike at the time of reset and (2) close to negative equity at the time of reset (a combined cumulative loan-to-value ratio of at least ninety). Borrowers resetting under these conditions face both a payment shock in meeting the higher loan payment subsequent to reset, and, due to their lack of equity, a payment constraint that is more likely to be binding, since refinancing is more difficult for borrowers facing this amount of leverage.

First, I categorize counties and zip codes along an index by the quarterly count of private-label mortgages facing a double-trigger reset. Over my sample, there is sizable variation in local geographies facing resetting ARM shocks, driven by a combination of differential originations in these product categories; different initial teaser length durations resulting in changes in timing of when initial resets occur; different interest rates inducing mortgages to reset under conditions of high interest rates; differences in leverage choices of households; and pre-reset changes in house prices, which also affect the loan-to-value calculation.

If foreclosures have important aggregate consequences, geographic areas experiencing a high exposure to double-trigger resets should experience greater foreclosures resulting from these resets, which may also impact local house prices and foreclosures among non-resetting mortgages in local areas. The presence of externalities resulting from mortgage default acts as an amplification mechanism of the initial reset shock inducing more foreclosures in areas.

To test this hypothesis, I regress the quarterly count of resetting ARMs against aggregate counts of foreclosures and price changes in the following quarter. If the external effects of foreclosures are large, I expect areas with a greater intensity of double-trigger default to

experience more foreclosures and greater price declines in subsequent quarters. I run the following specification:

$$O_{it} = \mu_{it} + \beta \cdot DT_{i,t-1} + \varepsilon_{it}$$

Where i indexes either zip codes or counties, and t indicates the quarter of observation. O_{it} is the outcome variable in question—either aggregate foreclosures or the local price changes. $DT_{i,t-1}$ captures the count of local double-trigger ARM resets, normalized by subtracting the mean and dividing by the standard deviation. The interpretation of the β coefficient measures the impact a one standard deviation rise in double-trigger resets in a quarter has on price changes and aggregate foreclosures the following quarter. To account for other changes in the economic environment, I include dummies for quarter of observation and fixed effects for the zip code or county. That is, I am not exploiting the fact that certain areas simply originated a greater fraction of ARMs than others—I use only the variation in *timing* of those resets and their severity as captured by resets that face upward interest rate shocks in conditions of negative equity. Standard errors are clustered at either the county or zip-code level.

Table IX shows evidence consistent with the hypothesis that foreclosure contagion has macroeconomic effects observable at the zip-code or county level. A one standard deviation rise in the relative frequency of double-trigger resets results in around a 0.3% drop in house measured at the zip-code or county level, relative to a typical mean of around a 0.6% drop. At the zip-code level, a one standard deviation rise in double-trigger resets leads to eleven additional foreclosures the following quarter (relative to an average of twenty-seven), or an additional 552 foreclosures at the county level (relative to a mean of 1330).

These effects are quite economically sizable and suggest a substantial role for foreclosure externalities in amplifying the effects of an initial set of foreclosures into larger economic losses for geographies affected by a greater than average number of adverse ARM resets. To be sure, this analysis lacks the tight identification of the microeconomic approach above, which relies on within-month variation in interest rates paid upon the time of reset. While I am able to control for a range of economic conditions through time and geography fixed effects, areas experiencing a wave of adverse reset shocks may face other simultaneous negative economic shocks. As such, these results should be seen as a descriptive illustration of the potential magnitude of the size of the aggregate consequences of foreclosure spillovers.

7 Conclusion

In this paper, I shed light on one important mechanism behind the massive wave of foreclosures experienced during the recent financial crisis: the role of default spillovers and peer effects. In understanding the externalities of foreclosures, I analyze two novel instruments influencing foreclosures: the choices of lookback period and financial index. I argue that these commonly overlooked features of ARM contracts allow for a quasi-experimental setting in which mortgages resetting in the same month are shocked at different interest rates for reasons that are highly unlikely *ex ante* to be closely related to default decisions of neighbors.

Consistent with prior literature on ARM reset, I first establish that the within-month variation in interest rates among resetting ARMs predicts default and foreclosure behavior among resetting loans. Next, I use the resulting spike in foreclosure rates among loans facing relatively higher interest rates (which are fixed for a period of six to twelve months) as a local shock to the foreclosure intensity experienced by their geographical neighbors. I find compelling evidence for the existence of foreclosure spillover effects: each foreclosure causes an additional 0.3–0.5 foreclosures in the 0.10 mile radius around a foreclosed property. These effects are economically large and point to large peer impacts of foreclosures.

While prior literature has examined price consequences of foreclosures, I find that it appears unlikely that price effects can fully explain the magnitude of my result. In addition to the price channel, I emphasize the role of a refinancing channel—I find that refinancing activity drops by a third in areas around foreclosures—as well as a neighboring default peer channel. I emphasize the role of informational and learning features of nearby mortgages. The presence of a nearby foreclosure provides local residents with important information both on the quality of the neighborhood as well as the viability of mortgage default. Consistent with this interpretation, I document that foreclosure spillovers are (1) concentrated in a tight geography around foreclosed properties, (2) highest among underwater properties, (3) do not manifest in defaults on non-mortgage debts, and (4) are highest in areas that have previously not experienced a foreclosure. This pattern is most consistent with an informational view of default in which borrowers strategically default in reaction to local foreclosures after learning about the cost of default in response to a visible nearby foreclosure.

I also document regional consequences of mortgage defaults. I focus on ARMs that reset under adverse circumstances—negative equity and upward rate resets—that are strongly predictive of future delinquency. In the presence of spillover effects, these defaults should impact borrowers in these areas more broadly. I document that zip codes and counties with a greater prevalence of adverse mortgage resets experience greater price drops and more foreclosures in the following quarter. These results provide complementary evidence on mortgage resets

in connection with prior work, which has documented the positive impacts of rate *decreases* among resetting ARMs on mortgage performance and local economic outcomes. While I am primarily interested in mortgage resets as an instrument to understand mortgage resets in general, my results also highlight the nature of interest rate-linked mortgage contracts as a key aspect of the monetary policy transmission mechanism.

Overall, these results highlight an important amplification channel associated with lender and borrower responses to defaults that resemble a traditional bank run. ARMs in particular can be seen as a form of repeated short-term borrowing, as borrowers typically refinanced at the end of their teaser period (at which point prepayment penalties frequently expired and the dynamics of rate reset could push interest rates higher). Continued repayment rates among loans featuring these contract details was predicated on the availability of continued access to market credit. When ARMs reset under conditions of low prices, however, refinancing was a challenge and borrowers faced with payment shocks often defaulted. To the extent that these defaults triggered price drops and falls in refinancing activity among neighboring homes, they led to a cascading wave of additional defaults and further price drops in surrounding areas. In the aggregate, the dynamics of peer interactions in the residential mortgage market help explain how an initially small shock to the subprime mortgage market was transmitted to a much greater shock among households broadly, and potentially points to important roles for public policy aimed at addressing foreclosures.

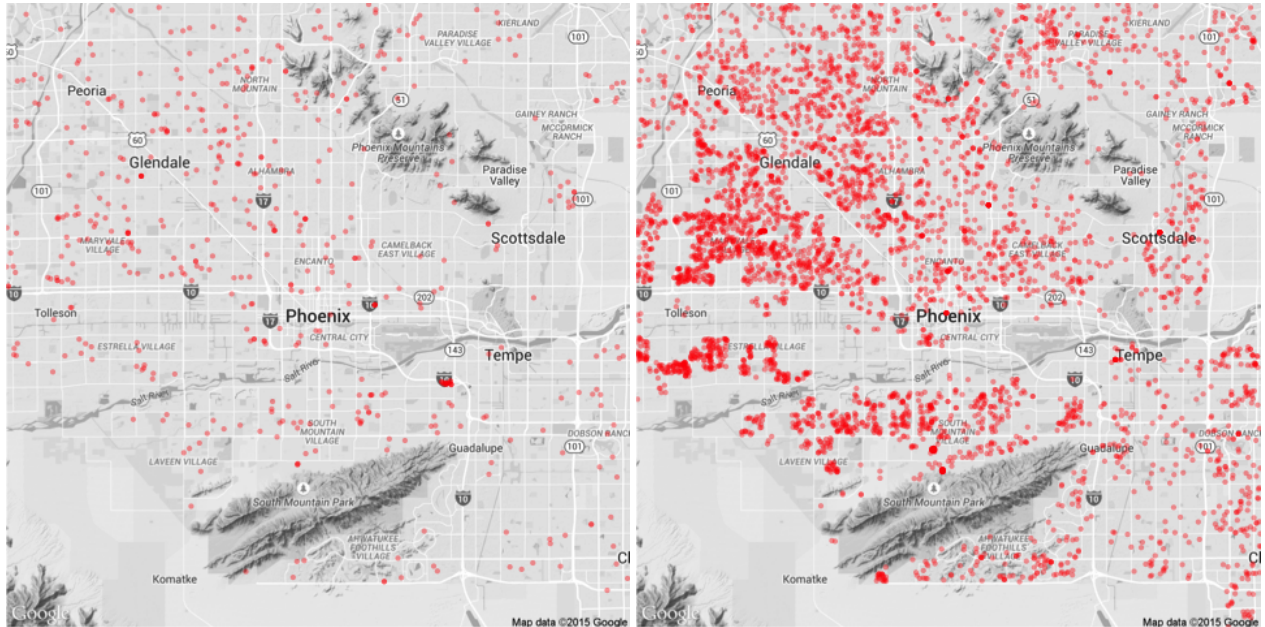
References

- Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, Tomasz Piskorski, and Amit Seru. 2012. "Policy Intervention in Debt Renegotiation: Evidence from the Home Affordable Modification Program." NBER Working Paper No. 18311.
- Agarwal, Sumit, Brent Ambrose, and Yildiray Yildirim. 2015a. "The Subprime Virus". *Real Estate Economics* 43: 891–915.
- Agarwal, Sumit, Gene Amromin, Souphala Chomsisengphet, Tomasz Piskorski, Amit Seru, and Vincent Yao. 2015b. "Mortgage Refinancing, Consumer Spending, and Competition: Evidence from the Home Affordable Refinancing Program. NBER Working Paper No. 21512.
- Anenberg, Elliot, and Edward Kung. 2014. "Estimates of the Size and Source of Price Declines Due to Nearby Foreclosures." *American Economic Review* 104(8):2527–51.
- Auclert, Adrien. 2015. "Monetary Policy and the Redistribution Channel." Working Paper.
- Azizpour, Shahriar, Kay Giesecke, and Gustavo Schwenkler. 2015. "Exploring the Sources of Default Clustering." Working paper.
- Benmelech, Efraim, Nittai Bergman, Anna Milanez, and Vladimir Mukharlyamov. 2014. "The Agglomeration of Bankruptcy." NBER Working Paper No. 20254.
- Bertrand, Marianne, Erzo Luttmer, and Sendhil Mullainathan, 2000. "Network Effects And Welfare Cultures," *The Quarterly Journal of Economics* 115(3):1019–1055.
- Blomquist, Daren. 2012. "2012 Foreclosure Market Outlook." RealtyTrac.com, February 13. <http://www.realtytrac.com/content/news-and-opinion/slideshow-2012-foreclosure-market-outlook-7021>.
- Breza, Emily. 2014. "Peer Effects and Loan Repayment: Evidence from the Krishna Default Crisis." Working paper. Columbia Business School.
- Bucks, Brian, and Karen Pence. 2008. "Do Borrowers Know their Mortgage Terms?" *Journal of Urban Economics* 64(2):218–233.
- Campbell, John Y., Stefano Giglio, and Parag Pathak. 2011. "Forced sales and house prices." *American Economic Review* 101(5):2108–31.
- Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." NBER Working Paper No. 19843.
- Chinco, Alexander and Christopher Mayer. 2015. "Misinformed Speculators and Mispricing in the Housing Market." Working paper.
- Dahl, Gordon, Katrine Loken, and Magne Mogstad. 2014. "Peer Effects in Program Participation," *American Economic Review*, 104(7):2049–74.
- Di Maggio, Marco, Amir Kermani, and Rodney Ramcharan. 2014. "Monetary Policy Pass-Through: Household Consumption and Voluntary Deleveraging." Columbia Business School Research Paper No. 14–24
- Fay, Scott, Erik Hurst, and Michelle J. White. 2002. "The Household Bankruptcy Decision." *American Economic Review* 92(3): 706–718.
- Fuster, Andreas and Paul Willen. 2015. "Payment Size, Negative Equity, and Mortgage Default." *Federal Reserve Bank of New York Staff Reports*, No. 582.
- Foote, Christopher, Kristopher Gerardi, and Paul Willen. "Why Did So Many People Make So Many Ex-Post Bad Decisions? The Causes of the Foreclosure Crisis." FRB Boston

- Public Policy Discussion Paper Series, No. 12–2.
- Fisher, Lynn M., Lauren Lambie-Hanson, and Paul S. Willen. 2013. “The Role of Proximity in Foreclosure Externalities: Evidence from Condominiums.” Public Policy Discussion Paper No. 13–2. Federal Reserve Bank of Boston.
- Gerardi, Kristopher, Eric Rosenblatt, Paul S. Willen, and Vincent Yao. 2012. “Foreclosure Externalities: Some New Evidence.” Working Paper 2012–11. Federal Reserve Bank of Atlanta.
- Glaeser, Edward, Bruce Sacerdote, and Jose Scheinkman. 2003. “The Social Multiplier.” *Journal of the European Economic Association*, 1(2-3):345–353.
- Goodstein, Ryan, Paul Hanouna, Carlos Ramirez, and Christof W. Stahel. 2011. GMU Working Paper in Economics No. 11–12
- Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2013. “The Determinants of Attitudes toward Strategic Default on Mortgages,” *The Journal of Finance*. 68:1473–515.
- Hull, John C. 2010. *Fundamentals of Futures and Options Markets*. 7/E. New York: Prentice Hall.
- Keys, Benjamin, Tomasz Piskorski, Amit Seru, and Vincent Yao. 2014. “Mortgage Rates, Household Balance Sheets, and the Real Economy.” NBER Working Paper No. 20561.
- Immergluck, Dan, and Geoff Smith. 2005. “There Goes the Neighborhood: The Effect of Single-Family Mortgage Foreclosures on Property Values.” Unpublished paper. Woodstock Institute, Chicago, IL.
- – –. 2006a. “The Impact of Single-Family Mortgage Foreclosures on Neighborhood Crime.” *Housing Studies* 21:851–66.
- – –. 2006b. “The External Costs of Foreclosure: The Impact of Single-Family Mortgage Foreclosures on Property Values.” *Housing Policy Debate* 17(1):57–79.
- Lusardi, Annamaria, and Peter Tufano. 2008. “Debt Literacy, Financial Experience, and Overindebtedness.” Unpublished manuscript. Harvard Business School.
- Munroe, David and Laurence Wilse-Samson. 2013. “Foreclosure Contagion: Measurement and Mechanisms.” Working Paper.
- Manski, Charles F. 1993. “Identification of Endogenous Social Effects: The Reflection Problem.” *The Review of Economic Studies* 60(3):531–42.
- Mayer, Christopher, Karen Pence, and Shane M. Sherlund. 2009. “The Rise in Mortgage Defaults.” *Journal of Economic Perspectives* 23(1):27–50.
- Mian, Atif R., Amir Sufi, and Francesco Trebbi. 2015. “Foreclosures, House Prices, and the Real Economy.” *Journal of Finance* 70(6):2587–2634.
- Palmer, Christopher. 2015. “Why Did So Many Subprime Borrowers Default During the Crisis: Loose Credit or Plummeting Prices?” Working paper.
- Parker, Jonathan A., Nicholas S. Souleles, David S. Johnson, and Robert McClelland. 2013. “Consumer Spending and the Economic Stimulus Payments of 2008.” *American Economic Review* 103(6):2530–53.
- Piskorski, Tomasz, Amit Seru, and James Within. 2015 “Asset Quality Misrepresentation by Financial Intermediaries: Evidence from the RMBS Market.” *Journal of Finance* 70(6):2635–2678.
- Smith, David Andrew, Louis Perwien, and Janneke Ratcliffe. 2009. “Mortgage Servicer Response to Borrowers in Crisis: A Report from the Front Lines.” Center for Community Capital. Working Paper.

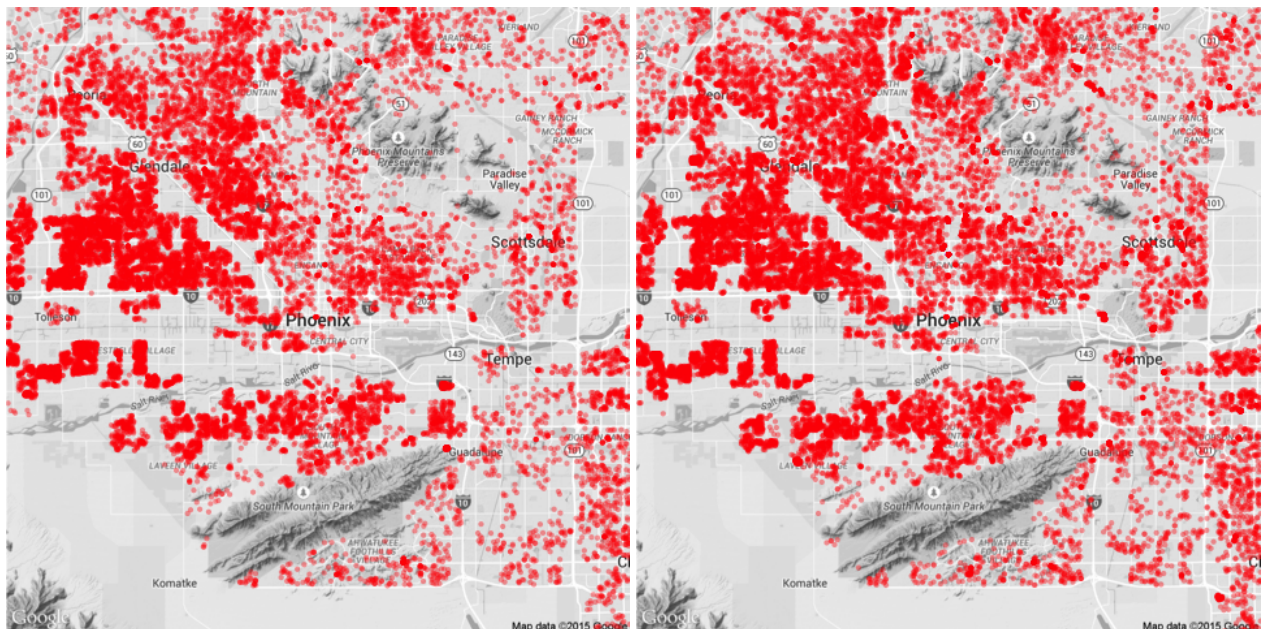
- Towe, Charles, and Chad Lawley. 2013. "The Contagion Effect of Neighboring Foreclosures." *American Economic Journal: Economic Policy* 5:313–35.
- United States. Cong. House Financial Services Committee. Hearing on Improving Foreclosure Prevention and Enhancing Enforcement. Dec. 6, 2007. 110th Cong. 1st sess. Washington: GPO, 2007 (statement of Sheila C. Bair, Chairman, Federal Deposit Insurance Corporation).

Figure I
Plots of Completed Foreclosures in Phoenix, AZ



2006

2007

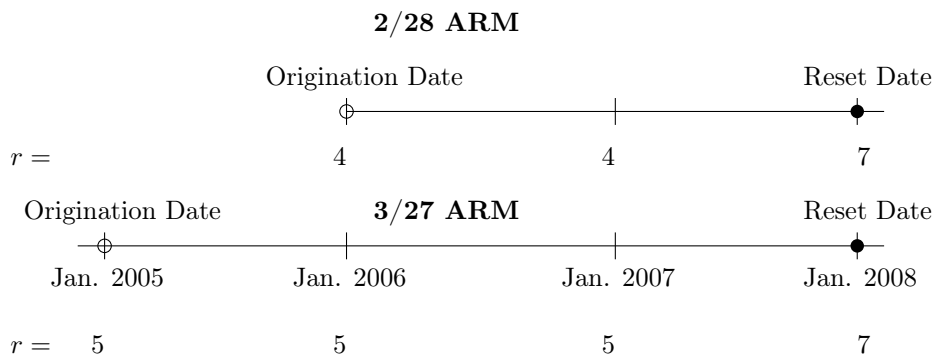


2008

2009

Figure II
Illustration of Reset Dates

Typical interest rate schedules are illustrated for two common mortgage products—A 2/28 ARM with an initial teaser period of two years, and a 3/27 ARM with an initial teaser rate of three years. Assuming both mortgages have identical margins and lookback periods, they pay identical interest rates subsequent to the reset.



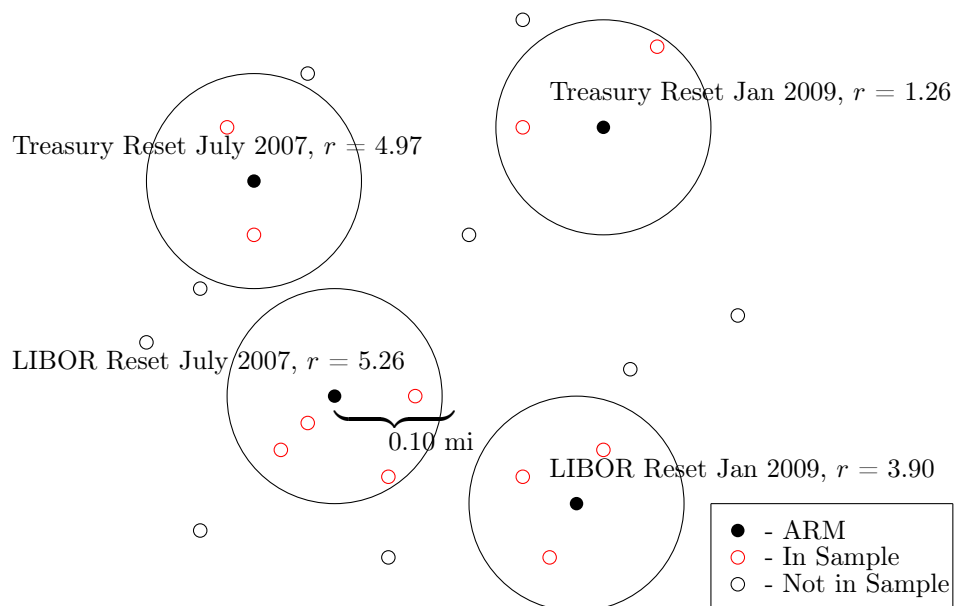


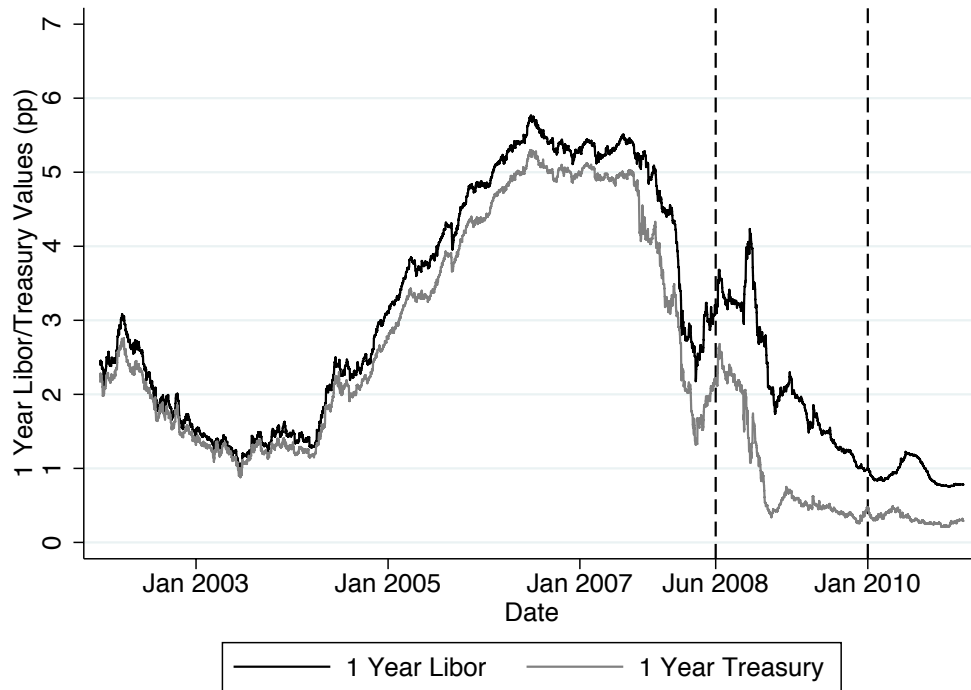
Figure III
Illustration of Sample Creation

This graph illustrates the mechanics of sample selection. Resetting adjustable-rate mortgages, represented as filled black circles, comprise the basis of the sample. Properties represented as red circles are included in the sample if they lie within a 0.10 mi radius of these loans. Properties represented by black circles are not within 0.10 miles of any resetting ARM are not included in the sample. The highlighted interest rates also illustrate the nature of the identification strategy—here, two sets of mortgages are resetting in the same month to different interest rates, exposing their neighborhoods to different foreclosure intensities.

Figure IV
Index Values

Panel A illustrates prevailing index rates on one year LIBOR and Treasury indices to which many adjustable-rate contracts were linked. Panel B illustrates the spread between these two indices. The spread reaches a maximum of over 300 basis points in late 2008. The areas highlighted by dashes indicate a period of large differences in interest rates between the two indices that are explored further in the next table.

Panel A: LIBOR and Treasury Values



Panel B: Spread between LIBOR and Treasury Indices

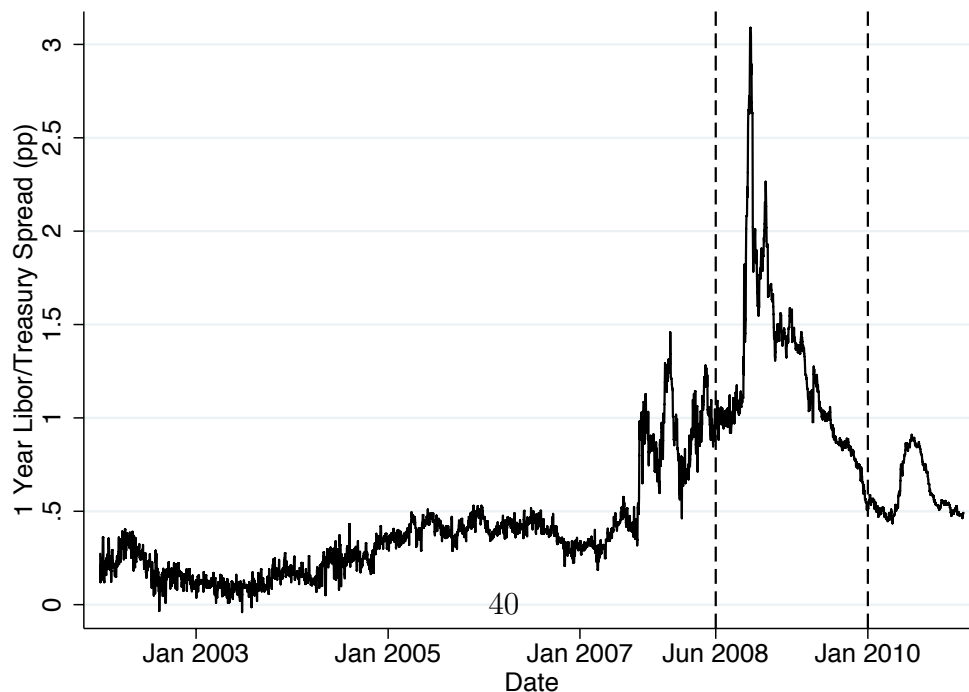
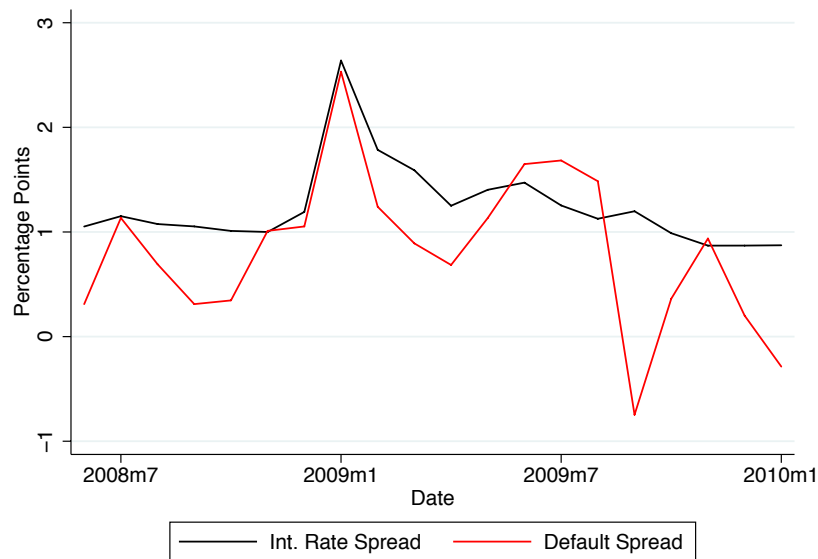


Figure V
Univariate Graphs: Interest Rate and Default Spreads

These graphs illustrate the differences in interest rates and default behavior among loans linked to different indices along with their neighbors. The black line in both panels illustrates the difference in interest rates paid among loans indexed to LIBOR relative to interest rates paid among comparable loans indexed to a Treasury index. This spread reaches a maximum in January 2009, when LIBOR-linked loans could expect to pay, on average, 260 basis points more than otherwise identical Treasury-linked loans—corresponding to about \$11,000 more in payments over the subsequent 12 months. Panel A illustrates how this interest rate differential corresponds to a difference in subsequent foreclosure rates among LIBOR-linked loans relative to Treasury-linked loans over the year following reset. Panel B illustrates how this interest rate differential translates into the probability of experiencing a foreclosure in the two years subsequent to the reset date among the *neighbors* of loans linked to either a LIBOR or Treasury index within a 0.10 mile radius.

Panel A: In-Sample LIBOR-Treasury Default Spread for Resetting Loan



Panel B: In-Sample LIBOR-Treasury Default Spread for Neighboring Loans

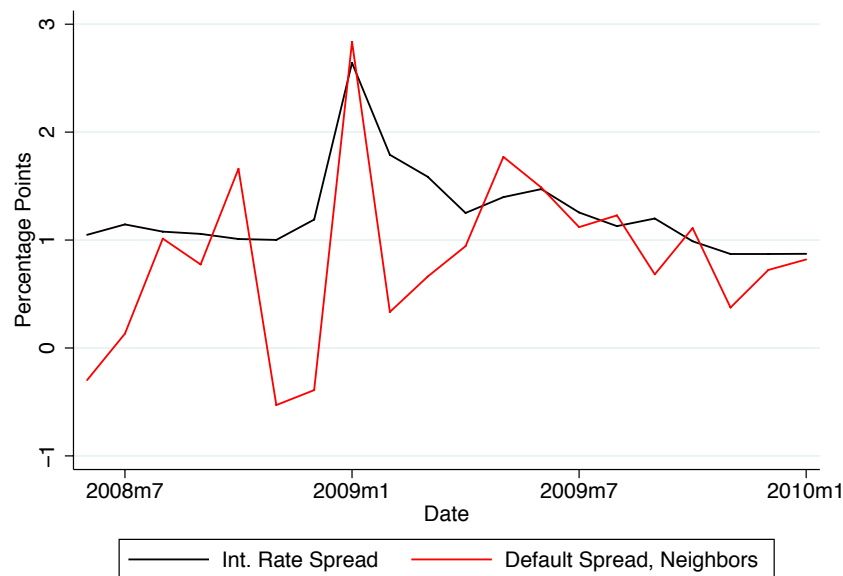


Figure VI
Illustration of Lookback Period

This graph illustrates the differences in potential interest rates paid by adjustable-rate mortgages resetting within the same month due to lookback variation. A mortgage with a lookback of 15 days will pay, for the next six or twelve months, an interest rate component of the payment drawn from the exact interest rate on the prevailing index (LIBOR or Treasury) 15 days prior; in this example 4.18%. An otherwise comparable mortgage with a lookback of 45 days will pay 5.02%.

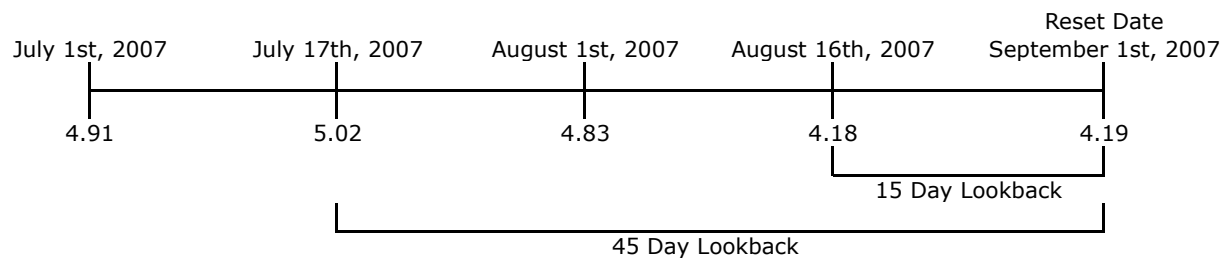
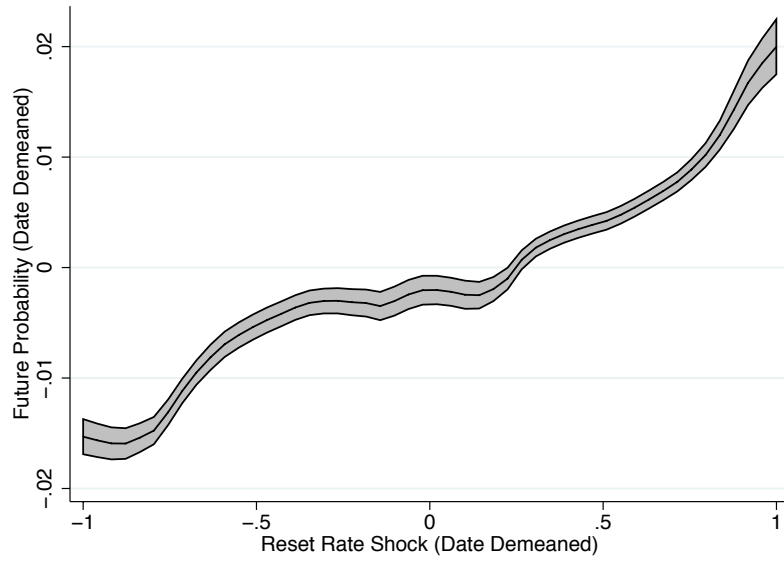


Figure VII
Default Responses to Interest Rate Hikes

These plots investigate the impact of within-month variation in reset rates on default behavior. Panel A illustrates a local polynomial smoothed plot of date-demeaned reset shocks against future date-demeaned foreclosure probabilities among resetting loans. Variation in reset rates, conditional on month, arises from differences in origination rates (due to differences in initial teaser term), index choice, and lookback period. Panel B uses the same measure of reset rate shocks among resetting adjustable-rate mortgages, but investigates the foreclosure responses among neighboring loans within a 0.10 mile radius in the subsequent two years.

Panel A: Impact of within-month variation in Interest Rates on ARM Default Rates



Panel B: Impact of within-month variation in Interest Rates on Neighbor Default Rates

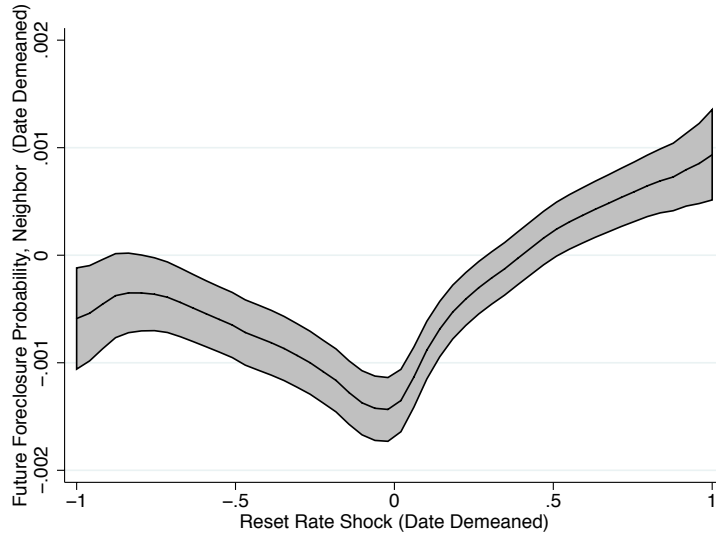


Table I
Summary Statistics: Lookback Sample

| Lookback Length: | 0 | 25 | 30 | 45 |
|---------------------------|------|------|------|------|
| <i>Resetting Loan:</i> | | | | |
| Credit Score | 723 | 704 | 714 | 704 |
| Original Balance (100\$k) | 266 | 259 | 263 | 262 |
| Conforming | 76% | 81% | 78% | 77% |
| CLTV | 87 | 107 | 99 | 92 |
| Refi | 49% | 48% | 49% | 49% |
| Investor | 9% | 11% | 7% | 8% |
| <i>Neighborhood:</i> | | | | |
| Density | 22 | 23 | 22 | 22 |
| Fraction Non-White | 36% | 36% | 35% | 37% |
| <i>Neighboring Loans:</i> | | | | |
| CLTV | 74% | 83% | 80% | 78% |
| Mortgage Amount | 207 | 241 | 218 | 213 |
| Year Built | 1967 | 1967 | 1967 | 1966 |
| ARM | 42% | 39% | 40% | 41% |
| Investor | 20% | 23% | 22% | 21% |
| N | 558k | 334k | 333k | 581k |

Table II
Summary Statistics: Index Sample

| Index Indicator: | LIBOR | Treasury |
|---------------------------|-------|----------|
| <i>Resetting Loan:</i> | | |
| Credit Score | 870 | 902 |
| Original Balance | 468 | 623 |
| Conforming | 31% | 9% |
| CLTV | 83% | 70% |
| Refi | 49% | 66% |
| Investor | 9% | 4% |
| <i>Neighborhood:</i> | | |
| Density | 24 | 22 |
| Fraction Non-white | 24% | 22% |
| <i>Neighboring Loans:</i> | | |
| CLTV | 77% | 67% |
| Mortgage Amount | 319% | 389% |
| Year Built | 1971 | 1968 |
| ARM | 42% | 48% |
| Investor | 23% | 20% |
| N | 395k | 287k |

Table III
OLS Evidence of Foreclosure Clustering

This table estimates the relationship between completed foreclosures and the observable number of additional foreclosure completions for different measures of radius lengths around a property through a benchmark OLS calculation. The sample consists of all loans within a 0.10 mi radius around a resetting adjustable-rate mortgage holder. The first column estimates the foreclosure response of borrowers in the two years following the mortgage reset; as a function of prior foreclosures in the entire radius prior to reset. The second column estimates the foreclosure response in the two years after reset among neighbors as a function of the foreclosure of the resetting adjustable-rate mortgage holder in the period after reset. Additional controls include a variety of variables taken at the level of both the resetting mortgage and the neighbors. Brackets contain the mean of the dependent variable for the sample in question. Standard errors are clustered at the tract-year level.

| | Local Foreclosures | Reset Foreclosures |
|-------------------------|--------------------------------|-------------------------------|
| Neighboring Foreclosure | 0.031*** (0.0005) [0.03] | 0.049*** (0.009) [0.03] |
| N(Clusters) | 4.5m (7468) | |

** denotes 5% significance, *** denotes 1% significance.

Table IV
Main Results

This table examines how within-month variation in interest rates, paid conditional on reset, predict future foreclosure behavior. The first stage regression is $D_{iz,t \rightarrow t+12} = \alpha + \mu_{zt} + \beta' X_{izt} + \gamma' R_{it} + \varepsilon_{izt}$, the reduced form (intention-to-treat) specification is $D_{-i,z,t \rightarrow t+24} = \alpha + \mu_{zt} + \beta' X_{-i,zt} + \gamma' R_{izt} + \varepsilon_{-i,zt}$, and 2SLS (instrumental variable) specification is $D_{-i,z,t \rightarrow t+24} = \alpha + \mu_{zt} + \beta' X_{-i,zt} + \delta' \hat{N}_{izt} + \varepsilon_{-i,zt}$; where the notation X_{-i} indicates that these values are taken among the set of all neighbors of the resetting ARM *except* the resetter herself. Brackets contain the mean of the dependent variable for the sample in question. Standard errors are clustered at the tract-year level. The dependent variable in the first column is the foreclosure propensity of the resetting mortgage in the year after reset; in the next two columns it is the foreclosure probability of neighboring properties for the two years after reset. The last column shows the F-test for the excluded instrument in the first stage of the regression.

| Dep Var: Foreclosure | First Stage | Reduced Form (ITT) | Second Stage (2SLS) | N(Clusters) | First Stage F |
|--------------------------------|--------------------------------|-----------------------------------|---------------------------------|-------------|---------------|
| Index Sample | 0.024*** (0.003) [0.025] | 0.0012** (0.0007) [0.02] | 0.050*** (0.013) [0.02] | 664k(14k) | 173 |
| Lookback | 0.025*** (0.002) [0.21] | 0.0005*** (0.00015) [0.041] | 0.016*** (0.005) [0.041] | 3.3m(45k) | 969 |
| All Resets | 0.025*** (0.0013) [0.17] | 0.0005*** (0.00014) [0.034] | 0.021*** (0.0049) [0.034] | 4.4m(58k) | 1031 |
| Resetter and Neighbor Controls | Yes | Yes | Yes | | |
| Fixed Effects | Zipcode | Zipcode | Zipcode | | |
| Dep Var is Foreclosure of: | Resetter | Neighbor | Neighbor | | |
| Key Regressor is: | Interest Rate | Interest Rate | Predicted Foreclosure | | |

** denotes 5% significance, *** denotes 1% significance.

Table V
Spillover Impacts on Refinancing Outcomes

This table examines how within-month variation in interest rates, paid conditional on reset, predict future refinancing behavior for neighbors around a resetting ARM for the following two years. Results are estimated both on a Deeds sample, for which refinancing activity is drawn from subsequent mortgage transactions, as well as on a BlackBox sample, for which refinancing is measured as the prepayment on an existing mortgage. The specifications shown reflect the instrumental variable specification in which the first stage regresses all within-month variation in interest rates against foreclosure rates of the resetting property, and the second stage uses the predicted value of foreclosure from these reset characteristics against refinancing among properties in the 0.10 mi radius around the resetting ARM. Standard errors are clustered at the tract-year level.

| | Refinancing (2SLS) | Refinancing (2SLS) | First Stage F |
|-----------------|-----------------------|-----------------------|---------------|
| Index Sample | | -0.078*** (0.027) | 1031 |
| Lookback Sample | | -0.062*** (0.099) | 969 |
| All Resets | -0.035*** (0.01) | -0.063*** (0.009) | 173 (458 BBX) |
| N(Clusters) | 1.7m(26k) | 4.4m(58k) | |
| Sample | BBX | Deeds | |
| Avg of Dep Var | [0.076] | [0.17] | |

** denotes 5% significance, *** denotes 1% significance.

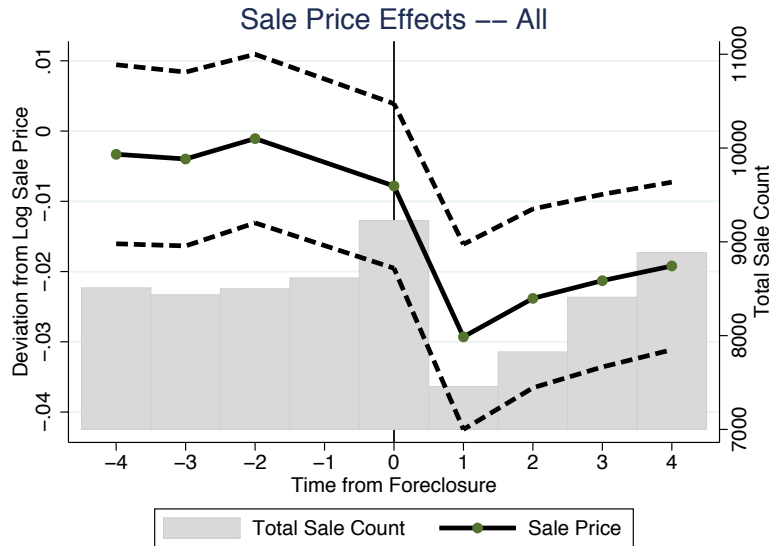
Figure VIII
Price effects

This graph illustrates the coefficients from a regression of log sale price among properties in the neighborhood of a resetting adjustable-rate mortgage against an indicator of timing relative to a completed foreclosure on the resetting loan. Time to foreclosure is measured in quarters relative to a completed foreclosure date on the adjustable-rate mortgage subsequent to reset. The estimates are drawn from a regression of the log sale price of properties against a variety of household characteristics:

$$\log(P_{it}) = \alpha + \mu_t + X_{it} + \sum_{k=-s}^{s-1} \mu_k \cdot 1[(t - T_i) = k] + \varepsilon_{it}$$

Where $\log(P_{it})$ is the log of the sale price of homes, fixed effects control for the month of observation, the zipcode, and a variety of property-specific controls. The key coefficients of interests are μ_k , which measure the change in sale price of the house during s quarters prior to and following the timing of the foreclosure in quarter T_i . Standard errors are clustered at the tract-year level. Panel A illustrates the response of all prices; Panel B illustrate the change in sale prices using a repeat-sale measure which incorporates the last selling price of the house. Background histograms capture total sale counts in the quarter in question.

Panel A: All Effects



Panel B: Repeat Sales

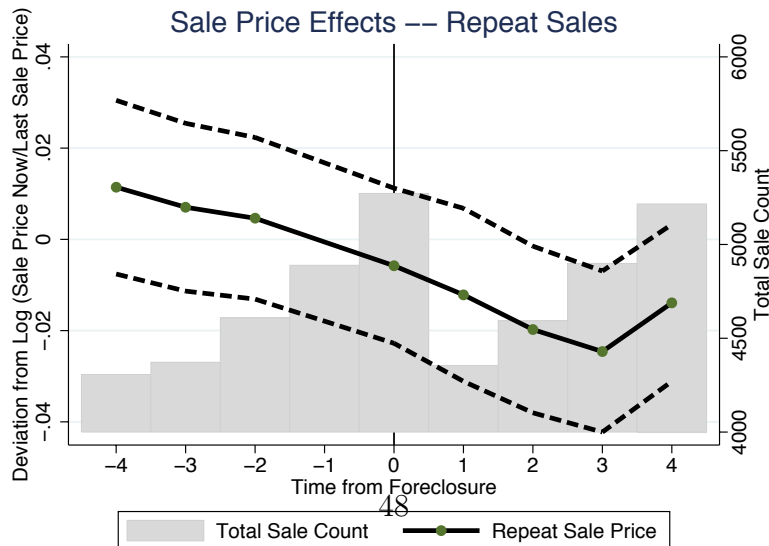


Table VI
Spillover Effects on Loan Default

This table examines how within-month variation in interest rates, paid conditional on reset, predict future default behavior for neighbors around a resetting ARM for the following two years. Results are estimated both on a Deeds sample, for which default indicators are taken from a notice of default field, as well as on a BlackBox sample, for which default is taken based on payment information of other private label securitized loans. The specifications shown reflect the instrumental variable specification in which the first stage regresses within-month variation in interest rates against foreclosure rates of the resetting property, and the second stage uses the predicted value of foreclosure from these reset characteristics against household default in the 0.10 mile radius around the resetting ARM. Standard errors are clustered at the tract-year level.

| | Notice of Default (2SLS) | Foreclosure Start (2SLS) | 90+ DPD (2SLS) |
|----------------|-----------------------------|-----------------------------|---------------------|
| All Resets | 0.019** (0.009) | 0.064*** (0.02) | 0.076*** (0.022) |
| N(Clusters) | 1.17m(26k) | 0.7m (54k) | 0.7m (54k) |
| First Stage F | 505 | 548 | 548 |
| Sample | Deeds | BBX | BBX |
| Avg of Dep Var | [0.076] | [0.19] | [0.25] |

Table VII
Heterogenous Effects

This table investigates how within-month variation in interest rates, conditional on reset, predicts foreclosure behavior. The specification is the same as in the third row and column of the main table IV, and uses the instrumental variable specification on all loans. Each row subsets on a different sample. The first four cuts subset on different characteristics of the neighbors; the last specification subsets on the equity position of the resetting mortgage. Properties are taken within a 0.10 mile radius of resetting ARM and standard errors are clustered at the tract-year level.

| Dep Var: Foreclosure | Second Stage (2SLS) Coefficient | N(Clusters) | First Stage F |
|----------------------|------------------------------------|-------------|---------------|
| Sample of: | | | |
| - FRM | 0.019*** (0.006) | 2.6m(56k) | 995 |
| - Apartment | 0.048 (0.015) | 0.5m (19k) | 112 |
| - Underwater | 0.03 (0.019) | 0.6m(44k) | 400 |
| - Investor | 0.024*** (0.009) | 0.9m(53k) | 320 |
| - Reseter Underwater | 0.023 (0.007) | 1.7m(32k) | 480 |

Table VIII
Effects by Prior Foreclosures

This table highlights the size of the foreclosure spillover effect by the intensity of previous foreclosures. Effect sizes are from an instrumental variable specification as illustrated as illustrated in the third column of Table IV among all resetters, broken out into different subsets based on the the number of previous foreclosures in a 0.25 mile radius around the resetting loan in the previous two years.

| Dep Var: Foreclosure | Second Stage (2SLS) | N | First Stage F |
|----------------------|------------------------|------------|---------------|
| Prior 0 | 0.019*** (.009) | 2.4m(27k) | 451 |
| Prior 1 | 0.023 (.012) | 0.7m (20k) | 198 |
| Prior 2 | 0.01 (0.018) | 0.4m(14k) | 76 |
| Prior 3 | 0.010 (0.021) | 0.3m(10k) | 51 |
| Prior 4 | -0.01 (0.026) | 0.2m(8k) | 52 |

Figure IX
Effects by Year

This graph illustrates the size of the spillover effects by year. Each point represents an instrumental variable regression as illustrated in the third row of the third column of column Table IV among all resettlers, broken out by the calendar year of reset.

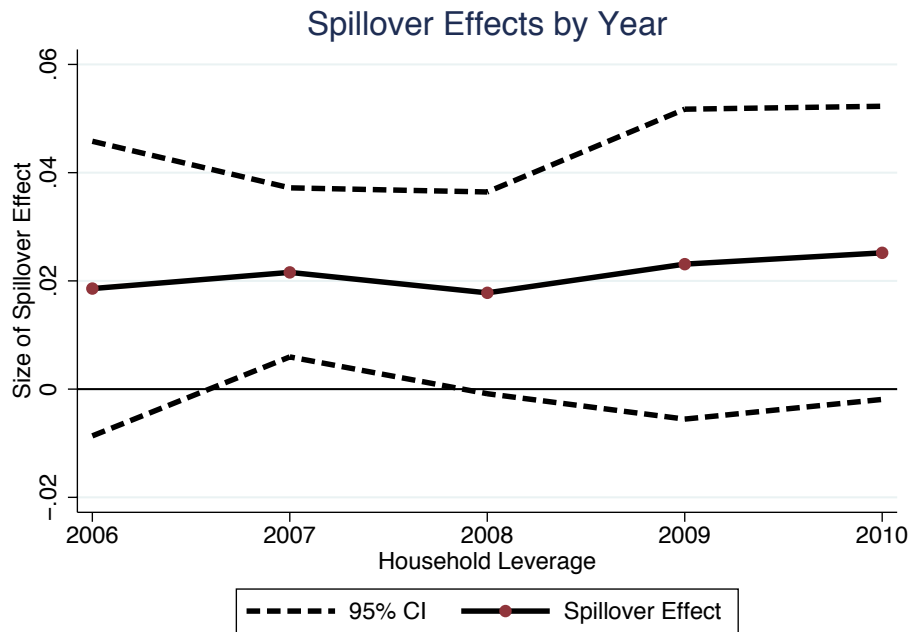


Figure X
Effects by Distance

This graph illustrates the size of the spillover effects by distance from the resetting loan. The sample includes BlackBox loans within 0.15 miles of a resetting property. Effect sizes are from an instrumental variable specification as illustrated as illustrated in the third column of Table IV among all resetters, broken out by different levels of distance from the resetting loan.

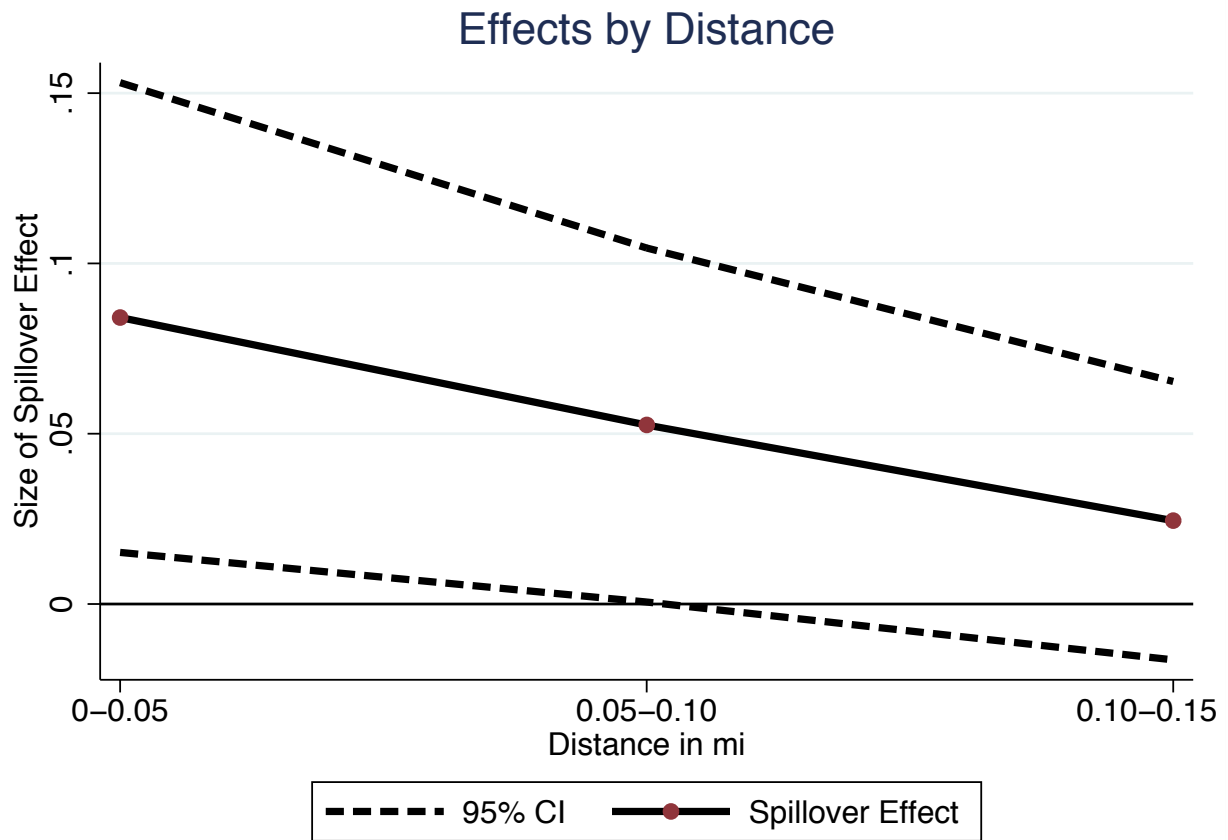


Table IX
Broader Macroeconomic Effects

This table illustrates the aggregate impact of mortgage resets. I identify “double-trigger” resets as those which face (1) An upward adjustment in interest rates at the time of reset, and (2) Conditions of near-negative equity (current combined loan to value of at least 90). Mortgage holders facing such conditions faced a combination of payment shock, difficulty in accessing financial markets for refinancing, and relatively high value of exercising their mortgage default option for strategic reasons—inducing these borrowers to default in greater proportion. I aggregate the number of such double trigger resetters in each quarter of observation among zip codes and counties, and regress a one standard deviation rise in double-trigger reset intensity against foreclosure and price impacts in the zip code or county the following quarter. I control for zipcode or county fixed effects, quarter of observation fixed effects, and cluster standard errors at the county or zipcode level.

| | Price | Foreclosure | Price | Foreclosure |
|----------------------------------|------------------------|----------------|-------------------------|-----------------|
| 1 SD in Double Trigger Resets | -0.0029*** (0.0008) | 552*** (73) | -0.0033*** (0.00016) | 11*** (0.61) |
| Level: | County | County | Zip | Zip |
| Avg of Dep Var: | -0.0064 | 1330 | -0.0067 | 27 |
| N(Clusters) | 746(54) | 1026(54) | 35k(2527) | 48k(2529) |

** denotes 5% significance, *** denotes 1% significance.

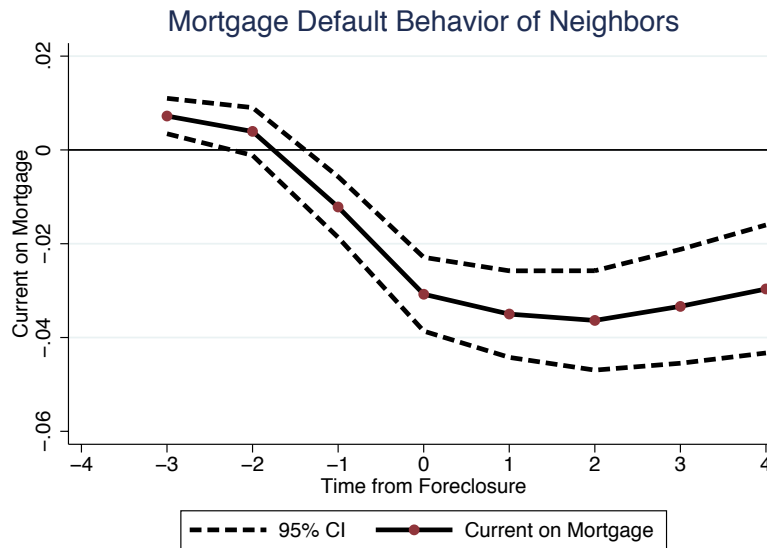
Figure XI
Timing of Neighbor Default Responses

This graph illustrates the coefficients from a regression of neighbor mortgage default behavior among mortgages in the neighborhood of a resetting adjustable-rate loan which experiences a foreclosure. Mortgage default information is taken from BlackBox; revolving debt default information is taken from Equifax and both are only included for private-label securitized loans in the neighborhood of the resetting loan. I run the regression:

$$O_{it} = \alpha + \mu_t + X_{it} + \sum_{k=-s}^{s-1} \mu_k \cdot 1[(t - T_i) = k] \varepsilon_{it}$$

Where O_{it} is the outcome of making payments on either a mortgage or other revolving debt (i.e., credit cards), fixed effects control for the month of observation, the zipcode, and a variety of property-specific controls. The key coefficients of interests are μ_k , which measure the change in sale price of the house during s quarters prior to and following the timing of the foreclosure in quarter T_i . Standard errors are clustered at the tract-year level. Panel A illustrates the propensity to make payments on mortgages as a function of time from the foreclosure of the resetting loan. Panel B illustrates payment behavior on revolving debt as a function of time from the foreclosure of a resetting adjustable-rate loan.

Panel A: Neighbor Responses on Mortgage Debt



Panel B: Neighbor Responses on Revolving Debt

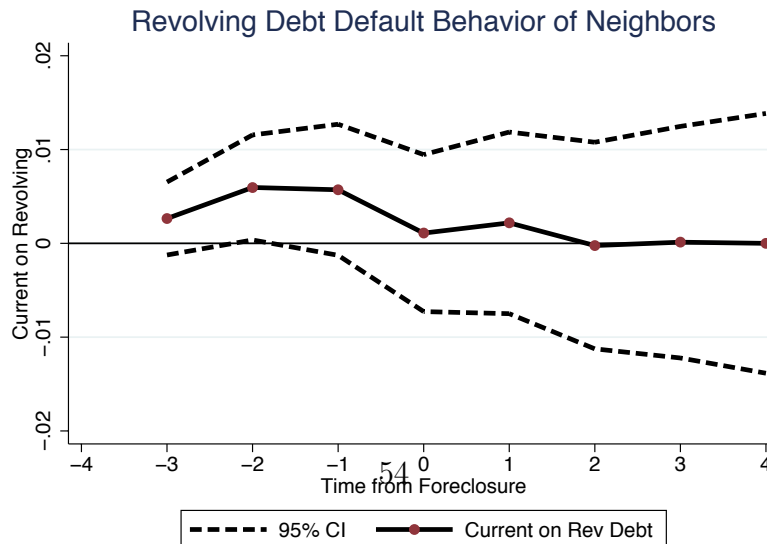


Table X
Placebo Test: Prior Foreclosures

This table performs a comparable instrumental variable analysis of reset characteristics on foreclosures as in the main analysis—but uses as a dependent variable in the second stage *prior* foreclosures among neighboring loans in the two years prior to reset. If the identification strategy is valid, within-month variation in interest rates should not predict foreclosures in the period prior to ARM reset.

| | BBX Sample | Deeds Sample |
|-------------|----------------------|---------------------|
| All Resets | -0.00007 (0.0027) | -0.0005 (0.0001) |
| N(Clusters) | 1.9m(57k) | 4.4m(58k) |

Appendix

Data Sources

Deeds Data The Deeds dataset is provided by DataQuick, a vendor which collects public-use transactions information. The data are organized at a property level and are comprehensive of all mortgage transactions which take place from 2000–2011 (foreclosure transactions typically go back further in time). The data list each mortgage transaction—including sales, transfers, new mortgages (first and second liens), and refinancing—which occur on a given property. I use the timing of the sales information to infer when cancer patients were resident in the property, and follow foreclosures for the duration of the time individuals were resident. I additionally use mortgage information dating to the time of the patient’s residence to calculate our key leverage statistics.

BlackBox Data BlackBox LLC is a private vendor which has collected the individual mortgage records related to private label securitized bonds (i.e., those not securitized by a government-sponsored entity like Fannie Mac or Freddie Mae). Though private label securitization made up only a fraction of total mortgage origination even at its peak before the crisis; the data contain more than 20 million mortgages in total; which is typically either subprime, Alt-A, or jumbo-prime in credit risk.

The BlackBox data contain static information taken at the time of origination, such as origination balance, credit score (FICO score), interest rate, and contract terms. The data are also updated monthly with dynamic information on fields like interest rates, mortgage payments, and mortgage balances. The mortgage payment field is most critical for our analysis, as it allows us to calculate the precise number of payments the household has made, not just whether or not the household has entered foreclosure.

Equifax Data Equifax is a major credit bureau which maintains detailed dynamic monthly credit information on households concerning their balances on mortgage and other debt, as well as credit scores (Vantage score).

Data Merges

A key innovation in this analysis is the use multiple sources of data on individual behavior to track financial outcomes around cancer diagnosis. This requires me to implement complex merges between many datasets which were not originally intended to be linked. Due to privacy restrictions, I am unable to make these data publicly available. However, below I document the merge process and linking variables which enable me to construct our dataset.

Deeds Data-BlackBox Though Deeds and BlackBox data were not designed to be linked, they are both administrative datasets containing reliable information on a variety of mortgage fields. I develop a novel a match method to link the two datasets using a training dataset (for which I know matches exactly) to develop the algorithm. The merge relies on the following common fields:

1. Exact date matches between origination dates of the mortgage are reported in the two datasets (not used if the origination date was likely imputed; i.e. the date reflected in BlackBox was the first or end of the month).
2. Zip code matches between the two datasets.
3. Matches based on mortgage purpose (i.e., refinancing or purchase).
4. Matches based on mortgage type (i.e., adjustable-rate or fixed-rate).
5. Matches based on mortgage origination amount (rounded down to the hundred)

I used a *backward* window of 31 days, in which the mortgage origination date reflected in BlackBox was at most 31 days after the date of the mortgage reflected in Deeds; and a *forward* window of 20 days.

The match algorithm worked by first focusing on (1) zip matches and (2) origination amount matches within the backward window (or the forward window if no matches existed in the backward window).

If only one match was found using those criteria, it was kept. If there were multiple matches, I restricted further by iteratively applying the following the following criteria. I first employed a “tight” match which required that the loan match uniquely on day, or (if there were multiple day matches) uniquely on mortgage purpose or type among those that matched on day.

If this did not uniquely identify a match, I next restricted to “looser” matches where there was (1) only one match uniquely on mortgage type and purpose. If no mortgage matched, I moved on to cases where there was (2) one unique match of either mortgage type or purpose with the other field missing, (3) one unique match on mortgage type, and (4) one unique match on mortgage purpose. The merge algorithm proceeded among all matching cases in the order specified above—if a high quality match was found, the mortgage was kept and the procedure only moved on to the other match cases in the order specified if no match was found.

BlackBox-Equifax BlackBox, a mortgage-level dataset, was linked by Equifax to borrower-level information on a variety of debts, including mortgages. The merge algorithm relied on a proprietary code which I cannot access. The vast majority of accounts in BlackBox were linked to a credit account.

To verify the accuracy of the merge, I imposed a restriction samples which make use of Equifax variables. Specifically, I require that the two entries match either on (1) zip code of the borrower (at least once over the life of the loan); or (2) have a match confidence of at least .85. The zip code restriction compares the zip code of the property as listed in BlackBox matches with the address of the borrower as listed in Equifax. A mismatched zip code is not necessarily indicative of a mismatch in loans—it could also suggest the presence of an investor who does not live in the property in question.

In addition to the zip code measure, Equifax provided a measure of match confidence ranging from 0–0.9. Loans at the top end of the confidence score reflect extremely well matched loans, and I allow for a mismatch in zip code so long as it is accompanied by a match confidence score of at I 0.85. Robustness checking based on other common attributes between the two datasets (such as common measures of default) suggest that the two measures of match accuracy I employ are effective in correctly identifying well-matched loans. For further details of the BlackBox-Equifax merge algorithm; see Piskorski, Seru, & Witkin (2015).

Data Cleaning

From the base Deeds data, the following cuts were made:

- Observations with a missing geocode, zipcode, or property id were dropped.
- I focused on a subset of 54 counties which accounted for about 40% of the relevant mortgages in my sample.

After linking the two datasets, I made the following restrictions to isolate loans for our analysis. I restricted to adjustable-rate mortgages with a recognizable reset date (ie, two, three, or five years after origination) and present indicator for the index type the loan is scheduled to reset to (LIBOR or treasury). I identified two samples: a 5/1 sample consisting of loans which reset yearly after five years to either a treasury or LIBOR rate, and a 2/28 + 3/27 sample consisting of loans which reset every six months after two or three years to a LIBOR rate. For both samples, I identified loans that reset appropriately as those loans which experienced their first interest rate change in a four month window around the scheduled reset date, and had no other interest rate resets within that window. I used the actual date of interest rate hike as reflecting the reset date (this coincided with the scheduled date change in the majority of cases). I further required that loans be current three months prior to this reset date. I focus on the first reset, and keep loan outcomes for the subsequent year.