Federal Reserve Bank of Philadelphia

Working Papers
Research Department

WP 20-37
September 2020
https://doi.org/10.21799/frbp.wp.2020.37

Mortgage Loss Severities: What Keeps Them So High?

Xudong An
Federal Reserve Bank of Philadelphia Supervision, Regulation, and Credit Department

Larry Cordell
Federal Reserve Bank of Philadelphia Supervision, Regulation, and Credit Department

Consumer Finance Institute

ISSN: 1962-5361

Mortgage Loss Severities: What Keeps Them So High?[*]

Xudong An
Federal Reserve Bank of Philadelphia

Larry Cordell
Federal Reserve Bank of Philadelphia

September 2020

Abstract

Mortgage loss-given-default (LGD) increased significantly when house prices plummeted during the financial crisis, but it has remained over 40 percent in recent years, despite a strong housing recovery. Our results indicate that the sustained high LGDs post-crisis is due to a combination of an overhang of crisis-era foreclosures and prolonged liquidation timelines, which have offset higher sales recoveries. Simulations show that cutting foreclosure timelines by one year would cause LGD to decrease by 5 to 8 percentage points, depending on the tradeoff between lower liquidation expenses and lower sales recoveries. Using difference-in-differences tests, we also find that recent consumer protection programs have extended foreclosure timelines and increased loss severities despite their potential benefits of increasing loan modifications and enhancing consumer protections.

Keywords: loss-given default (LGD), foreclosure timelines, regulatory changes, Heckman two-stage model, accelerated failure time model

JEL Classification Codes: G21, G18, C41, C24, G01

---

[*] Contacts: Xudong An, Federal Reserve Bank of Philadelphia, Ten Independence Mall, Philadelphia, PA 19106, xudong.an@phil.frb.org; Larry Cordell, larry.cordell@phil.frb.org. We thank three anonymous referees and Sumit Agarwal, Paul Calem, Brent Ambrose, Jamie Amico, Michael Bradley, Dennis Capozza, Phil Comeau, Shawn Connell, Andy Davidson, Ralph DeFranco, Hamilton Fout, Scott Frame, Liang Geng, Kris Gerardi, Doug Gordon, CL Kuo, Michael LaCour-Little, Lei Li, Brian McCarthy, Doug McManus, Hemachandra Mantrala, Deng Ning, Rushi Patel, Lan Shi, Bob Van Order, Paul Willen, Weifeng Wu, Peter Zorn, and seminar and conference participants at Fannie Mae, Freddie Mac, FHFA, the 2016 Inter-agency Risk Quantification Forum (OCC), and the 2017 AREUEA National Conference for helpful discussions and comments. We also thank Ruth Parker for excellent assistance in collecting mergers and acquisitions data on mortgage servicers, and both Ruth and Barb Brynko for editing this manuscript. The first draft of this paper was circulated under the title "Regime Shift and the Post-Crisis World of Mortgage Loss Severities" in December 2016. This article has been accepted for publication and has undergone full peer review. It has been through copyediting, which may reflect differences between this version and the version of record. Please cite this article as https://doi.org/10.1111/1540-6229.12334.

1. Introduction

Because of the existence of real property as collateral, mortgage lenders and investors usually recover a large portion of their loan in case of default and thus only incur partial losses. During the early 2000s, the average loss severity rate, or loss-given-default (LGD), for Fannie Mae and Freddie Mac loans was less than 20 percent. Loss severities on liquidated loans at the two government-sponsored enterprises (GSEs) more than doubled during the housing crisis, reaching a high of 48 percent in 2011. Despite a full recovery of the housing market, loss severities remained around 40 percent in recent years (Figure 1). Using recently released loss data from the GSEs, this paper is the first to directly observe that losses have remained persistently high post-crisis and link this fact to the increase in liquidation timelines that have been observed in prior studies (see, e.g., Cordell et al., 2015 and Cordell and Lambie-Hanson, 2016). In addition, we provide some causal evidence on the factors that contributed to the lengthening of liquidation timelines.

The GSE data we exploit are detailed loss information in Freddie Mac's Single-Family Loan-Level Dataset and Fannie Mae's Single-Family Loan Performance Data (hereafter Freddie Mac data and Fannie Mae data, respectively, and GSE data altogether). The loan-level datasets contain over 50 million loans, among which more than 1 million loans defaulted and were ultimately liquidated. The sheer size of the data, as well as the richness of its content, is unparalleled. For one, it breaks out losses into their components. The data also span the years from 1999 to 2017 and thus covers the full housing boom, bust and recovery period of the recent housing crisis, providing us with a unique opportunity to conduct in-depth analysis on both the cross-sectional and time-series variations in loss severities during a period of immense change in the mortgage market and regulatory environment. It also enables us to address several limitations in existing studies, such as sample selection bias and the inability to make causal inferences about market and regulatory developments.

We start by laying out the macroeconomic environment and the major loss severity components covered by the data. Our simple bivariate plots show that, during the crisis, high loss severity rates were associated with historic declines in house prices and an enormous pipeline of defaulted mortgages that likely strapped limited servicing resources. In recent years, sales recoveries have improved significantly, and defaults have declined. However, liquidation expenses rose from less than 6 percent in 2009 to more than 16 percent during 2014–2017; the soaring expenses are closely tied to lengthening timelines from foreclosure initiation to liquidation. It is also noteworthy that, among the loans that were liquidated, the share of risky, poorly underwritten mortgages originated during the 2005–2008 boom remained high in recent years, further contributing to high loss severities.

We then use the loan-level data to conduct multivariate analysis. Because there may be selection bias associated with selection from the full population of defaulted loans into the subset that proceeds through to liquidation, we run a Heckman two-stage model. Selection into liquidation is modeled in the first stage and the inverse Mills ratio is included in the second-stage loss severity regressions. Results show that loss severity is significantly and positively associated with the loan-to-value (LTV) ratio at liquidation. Liquidation timelines are shown to be a strong driver of loss severity — an extension of liquidation timelines by one year could be associated with as high as eight percentage points (or a 20 percent) increase in loss severities, ceteris paribus. We also find that the large lenders fined for fraudulent foreclosure practices and subject to new servicing rules under the National Mortgage Settlement (NMS) and the Consumer Financial Protection Bureau (CFPB) had higher loss severities.

Existing studies show that increases in liquidation timelines are associated with lower recoveries in liquidation sales, depending on the house price trajectory (see Cordell et al., 2015, and Goodman and Zhu, 2015). Therefore, to learn the true impact of the shortening/lengthening of the liquidation timelines, we run simulations based on the loss severity model estimated. Results show that the impact of cutting foreclosure timelines by one year increases recoveries from 5 percentage points to 8 percentage points, depending on where the property is located.

To gain insights into the drivers of prolonged liquidation timelines, we estimate Accelerated Failure Time models following Cordell et al., 2015. The results show that loans liquidated in recent years are associated with significantly longer timelines, even after controlling for local market conditions and loan and borrower characteristics. The existing literature points to recent mortgage market developments such as the new servicing rules by the CFPB as potential contributing factors (see, e.g., Cordell and Lambie-Hanson, 2016). To draw conclusions about causality, we conduct difference-in-differences (DID) tests and show that the NMS and the CFPB mortgage servicing rules have extended foreclosure timelines, ultimately leading to higher loss severities.

Our analysis contributes to the academic literature in several ways. First, in contrast to the abundant literature on default probability[1], research on LGD is sparse, mainly because of data limitations. In this paper, we add to the LGD literature by leveraging the newly available datasets from the GSEs. Existing LGD studies usually cover only a brief timeframe, and thus limit their ability to study LGD dynamics over time (see, Lekkas, Guigley, and Van Order, 1993; Berkovec, Canner, Gabriel, and Hannan, 1998; Calem and LaCour-Little, 2004; Capozza and Thomson, 2005; and Qi and Yang, 2009). Our current research covers the complete boom, bust, and recovery cycle surrounding the financial crisis. We directly observe that losses have remained persistently high post-crisis and link this fact to the increase in liquidation timelines that have been observed in prior studies.

Second, this is one of the first papers to study the breakout of the components of loss severity by exploiting the detailed loss information in the GSE data.[2] This turns out to be important for us to pinpoint the exact drivers of loss severity dynamics. Through the analysis, we are able to show that many of the liquidation expenses, which were counted as default loan losses are time-varying and thus higher liquidation expenses associated with longer liquidation timelines are determinants of higher loss severities. We are also able to demonstrate the tradeoff between higher sales recoveries during a housing market recovery and longer liquidation timelines (and thus expenses), raising interesting questions about liquidation timeline management and optimization.[3]

Third, most previous research on mortgage LGD do not consider potential sample selection bias associated with observing LGDs on only defaulted loans that were liquidated. To obtain unbiased estimates of loss severity parameters and to build an LGD model that can be used to forecast and price credit risk of all loans, it is important to correct for sample selection bias. We do this through a Heckman two-stage model.

[1] See, e.g., Campbell and Dietrich (1983); Deng, Quigley, and Van Order (2000); Clap, Deng, and An (2006); Demyanyk and Van Hemert (2011); Mian and Sufi (2009); Keys et al. (2010); Elul et al. (2010); An, et al. (2012); Agarwal et al. (2014a); Rajan, Seru, and Vig (2015); and An, Deng, and Gabriel (2019).

[2] We recently learned that Le and Pennington-Cross (2019) also used the Freddie Mac data to study loss on sale and holding costs.

[3] Le and Pennington-Cross (2019) also studies the breakout of loss severities using Freddie Mac data.

Fourth, we provide causal evidence on foreclosure timelines of the impact of recent consumer protection programs, namely the National Mortgage Settlement (NMS) of 2012 and the CFPB mortgage servicing rules starting in 2014. The new servicing rules associated with these programs were intended to enable more borrowers who are experiencing repayment difficulties to remain in their homes through increased loan modifications. Cordell and Lambie-Hanson (2016) pointed to this effect and possible adverse effect on foreclosure timelines. We provide causal evidence that these rules have lengthened foreclosure timelines and elevated foreclosure costs. Thus, we shed light on a potential cost-benefit tradeoffs of the new servicing rules meant to protect consumers.

Finally, the loss severity and liquidation timeline models we develop, along with our analysis of the factors contributing to rising loss severities, can help the investment community more accurately estimate and price mortgage credit risk associated with mortgage-backed securities and the GSEs' fast-growing credit risk transfer (CRT) deals.[4] The analysis can also provide useful input to bank stress testing exercises and to address new accounting regulations governing financial institutions' reserving practices, namely the Current Expected Credit Loss (CECL) requirements.[5]

The rest of the paper proceeds as follows. In Section 2, we describe our data and bivariate analysis on loss severity trends. In Section 3, we present our multivariate analysis of loss severity components and the Heckman two-stage model results. We explain our liquidation timeline model and the difference-in-differences tests in Section 4. Conclusions and areas for future work are discussed in Section 5.

[4] According to the most recent analysis, Layton (2020) estimates that CRTs now provide credit protection on some 70 percent of the $5 trillion of mortgage loans in GSE securities, $3.5 trillion worth. In these deals, the GSEs give up some of their guarantee fee to transfer some of their default risk to private investors. In the earliest deals, LGDs were set at a fixed percentage, but since then, bond support levels are based on total losses, requiring investors and rating agencies to estimate total losses, one reason for the enhanced disclosures by the GSEs.

[5] As discussed in deRitis and Tudor (2016), Current Expected Credit Loss (CECL) rules are a new requirement from the Financial Accounting Standards Board (FASB) starting in 2020 requiring banks to hold reserves based on a life-of-loss (LOL) concept, replacing the current impaired loss method. Models such as ours are applicable to this modeling problem.

2. Data

To promote their credit risk transfer (CRT) deals, starting in 2014, the GSEs began releasing to the public detailed loan-level performance data on their first-lien, full-documentation, fixed-rate mortgage (FRM) loans originated starting in 1999. We obtained these data directly from GSE websites.[6] We also obtained supplemental data to help in our modeling, including market interest rates; the CoreLogic Solutions (hereafter CoreLogic) house price index (HPI); mortgage performance information from a third source, Black Knight McDash (hereafter McDash); and servicer merger and acquisition (M&A) data from various online sources. They are all matched with our mortgage data for this study.

The full credit performance dataset released by the GSEs contains more than 50 million loans. Not all loans went into default, and among defaulted loans, only a subset were liquidated. Therefore, loss information is available for a fraction of those 50 million loans that were liquidated through property disposition, either through a conventional real estate owned (REO) disposition or one of the foreclosure alternatives, such as a short sales, foreclosure sales, note sales, or full charge offs. We describe these different liquidation types in Appendix Figure 1. Our final mortgage loss sample contains 898,220 loans that were originated between 1999 and 2016 and liquidated from 2000 to 2017.[7] The loss information includes last interest paid date, loan liquidation date, type of liquidation, REO date (if REO), default unpaid principal balance (UPB), liquidation expenses, net sale proceeds,[8] recoveries from mortgage insurance (MI), and other credit enhancements, as well as non-MI recoveries.[9]

In addition to loan loss information, GSE data provide detailed information on loan characteristics, including original and current principal balance, note rate, LTV ratio and combined LTV at origination,[10] borrower origination credit score, loan purpose, property occupancy status, MI

[6] See http://www.fanniemae.com/portal/funding-the-market/data/loan-performance-data.html and http://www.freddiemac.com/news/finance/sf_loanlevel_dataset.html.

[7] We exclude repurchase loans and loans with missing values for critical information such as last interest paid date. We also exclude loans reported to have losses "covered" but have no detailed loss information.

[8] Selling expenses are subtracted from gross sales proceeds to derive net sale proceeds.

[9] The GSEs provide a detailed breakdown of expenses such as legal costs, property maintenance and preservation costs, and tax and insurance payments.

[10] Combined LTV includes all liens at origination. It does not include liens recorded after origination.

coverage percentage, the three-digit property zip code,[11] the name of the servicer, and the name of the mortgage seller.

The GSE data also contain hundreds of millions of monthly performance records on all (defaulted and non-defaulted) loans, including their payment and delinquency status and the ultimate disposition type (voluntary repayment, REO, or foreclosure alternative).[12] This event-history data allow us to identify the whole default timeline, calculate the equity position of each liquidated loan through its entire life, and calculate other time-varying variables.

In Table 1, we provide summary statistics for the 898,220 liquidated loans as well as for the 52,122,741 loans (defaulted and non-defaulted) in our sample. For liquidated loans, the average loan balance at origination is about \$167,000, and the average default UPB is about \$158,000. The balance weighted average note rate for liquidated loans is 6.3 percent; the mean LTV ratios at origination and liquidation are 80 percent and 93 percent, respectively; and the average borrower credit score and DTI ratios are 639 and 39 percent, respectively.[13] It took an average of 19 months (from serious, or 90-day, delinquency) to liquidate those loans, and 69 percent of the liquidations were through REO sales. Not surprisingly, the full sample of loans, including defaulted and non-defaulted loans, have better risk profiles than liquidated loans. They have lower origination LTVs, higher borrower credit scores, and lower DTIs. They also carry lower interest rates.

We define the loss severity rate as the total net loss divided by the UPB at liquidation, in which total net loss includes net sales losses (net sale proceeds minus UPB), and liquidation expenses offset by non-mortgage insurance (non-MI) recoveries. We do not account for recoveries from MI or other credit enhancements as those recoveries to the GSEs are still losses (to others), which is the subject of our analysis. Carrying costs are not considered either.[14] Therefore,

---

[11] Because of privacy concerns, the GSEs do not provide the five-digit zip code. The SEC only provides a two-digit zip code as part of the new Regulation ABII asset-level disclosures.

[12] Freddie Mac lumps short sales, short payoffs, deeds-in-lieu of foreclosure, and foreclosure sales together.

[13] Hereafter, all mean and average numbers are balance weighted.

[14] There is no consensus on how carrying costs should be calculated. We believe carrying costs should be calculated in two parts: The first part is the interest cost in the first four months the servicer advances principal and interest to investors, which gets reimbursed to the servicer; the second part is GSEs' carrying cost after the loan is removed from the MBS pool and placed onto GSEs' balance sheet. GSEs indicate a calculation where

$$\text{loss severity rate} = \frac{UPB + expenses - sales - nonMIrecovery}{UPB}. \quad (1).$$

As shown in Table 1, loss severities of GSE loans average 42 percent (default UPB weighted). For some loans, loss severity is over 100 percent. This can happen if a lengthy process of liquidation results in significant carrying costs and liquidation expenses or total charge offs that include other expenses. As discussed previously, the average time from serious delinquency (SDQ) to liquidation is about 19 months. In fact, nearly 12 percent of the loans were liquidated three or more years after becoming seriously delinquent. Net sale proceeds are on average 67 percent of UPB. About 20 percent of the time, sales recoveries are less than 40 percent. Average liquidation expenses are about 10 percent of UPB; in some cases, they are as much as 44 percent. Non-MI recoveries average around 1.2 percent.

In Figure 1, we plot average loss severity rates of GSE loans by year of liquidation. In the same chart, we show the SDQ rate of conventional prime FRM loans based on the Mortgage Bankers Association's National Delinquency Survey.[15] From 2000 to 2003, loss severity rates are very low, less than 20 percent. Starting from 2005, severity rates rose significantly, from 23 percent in 2005 to 30 percent by 2007. The timing of this severity rate increase is interesting because the national housing market did not turn down significantly until 2007, with SDQ rates remaining low through 2007. This is not surprising when we consider that we observe loss severities only on loans that defaulted. Loans that went through default and liquidation before 2008 were most likely problem loans in areas that did not experience the house price gains in other areas.[16] The sharp rise in loss severities from 2007 to 2011 is clearly related to the sharp decline in house prices nationally during that period. Rising loss severities can also be attributed to the sharp rise in delinquent loans, shown by the dashed line in Figure 1. In Figure 2, we show liquidation volumes and shares of foreclosed properties by state. These rising foreclosures during the crisis overwhelmed mortgage servicing resources and court dockets in some judicial states, leading to increased severities.

---

mortgage note rate is used to calculate carrying costs during the full liquidation timeline. Our subsequent analysis findings are invariant to including carrying cost or not.

[15] We access the data through Haver Analytics.

[16] Michigan and Ohio were the two top states with defaulted loans in the mid-2000s; California was the top state after 2009.

What is most surprising about these figures is the high loss severities post-2012, a period of rising house prices combined with declines in delinquencies and the foreclosure pipeline. Loss severities remained over 40 percent from 2013–2016. In Appendix Figure 2, we separate loans by vintage and show a similar trend in each loan cohort. Across all vintages shown in the chart, the loss severity contours are similar to that of all vintages combined (the solid line) — basically a steady growth all the way to the recovery period. If the rationale for the high loss severity rates during the crisis was the historic declines in house prices and limited servicer/court capacity to manage surging delinquencies/foreclosures, the historically high loss severities that have persisted since 2012 are a puzzle as the house price and the liquidation capacity problems have abated. Empirically examining this puzzle is a focus of the remainder of this paper.

3. What Keeps Loss Severities So High?

3.1. Loss Severity Components and Their Trends

An advantage of analyzing drivers of the observed loss severity dynamics in the GSE data is that they provide detailed information on the components of loss severity. In the first three panels of Figure 3, we plot the loss components, including net sales recoveries, liquidation expenses, and non-mortgage insurance (non-MI) recoveries, all by year of liquidation. The cyclicality of net sales recoveries is clearly evidenced in Panel 1: The average sales recovery ratio was over 90 percent during 2001–2002; it fell below 60 percent in 2011, and then came back close to 80 percent in 2017. On the same graph, we plot the national Federal Housing Finance Agency (FHFA) home purchase house price index. Clearly the changes in sales recovery ratios coincide with the ups and downs of house prices surrounding the financial crisis. The correlation between the two series is 92 percent during 2005–2015, even though in the last two years sales recoveries have not kept up with house price appreciation.

What is striking is the increase in expenses in recent years, as evidenced in the second panel of Figure 3. Liquidation expense ratios rose to 12 percent in 2013 and 18 percent in 2015, and then remained at about 17 percent in 2016 and 2017. In contrast, they were as low as 5 percent in 2000 and 6 percent in 2009.

During the foreclosure process, servicers and GSEs incur various types of expenses, from loan workout trial expenses to tax and insurance payments to property maintenance and preservation costs. In fact, in the data, Freddie Mac breaks out liquidation expenses into Legal Costs, Maintenance and Preservation Costs, Taxes and Insurance, and Miscellaneous Expenses; and Fannie Mae breaks out liquidation expenses into Foreclosure Costs, Associated Taxes for Holding Property, Property Preservation and Repair Costs, Asset Recovery Costs, Miscellaneous Holding Expenses and Credits. Several expenses such as tax and insurance payments, maintenance and preservation costs, and administrative expenses are time related. In Panel 2 of Figure 3, we plot the average liquidation timelines, which we define as the number of months lapsed between serious (90-day) delinquency and final liquidation. The contours of expense ratio and timelines are almost identical. The correlation between the two series is 94 percent. Apparently, the surge in liquidation expenses explains a big part of the lingering high loss severities in recent years.

Non-MI recoveries include non-sale income such as tax and insurance refunds, hazard insurance proceeds, rental receipts, positive escrow, and other miscellaneous credits. As evidenced by Panel 3 of Figure 3, there are variations in non-MI recoveries over time. They are low during 2008–2010 and become higher in recent years. Even though the non-MI recoveries are a small portion of the recoveries (on average 1.2 percent), they tend to offset losses and are shown to have a negative relation with loss severities, especially in recent years.

It is well-known that the 2005–2008 vintages of loans were risky. In Panel 4 of Figure 3, we show that loss severities of the 2005–2008 vintages are significantly higher than those of other vintages. On the same chart, we plot the share of liquidations that are from the 2005–2008 vintages. It has been over 70 percent during 2009-2015, while it declined significantly in the last two years. This confirms the conventional wisdom that severities were driven up by the large volume of risky legacy loans whose liquidations were delayed because of various moratoria and the freezing of foreclosures during the "robo-signing" scandal.[17] They also appear to have contributed to the high loss severities we observe in recent years.

3.2. Heckman Two-Stage Loss Severity Model

We only observe loss information on loans that were liquidated, which is only a portion of the loans that defaulted. More important, the selection of loans into liquidation is likely nonrandom. Therefore, we need to account for potential sample selection bias in a loss severity regression to

[17] These events are described in the details that follow.

obtain unbiased estimates of loss severity parameters. We run a standard Heckman two-stage model for that purpose.

In the first-stage, we estimate the following Probit selection model:

$$Prob(L_i = 1|Z_i) = \Phi(Z_i\gamma), \quad i = 1, \dots, N, \tag{2}$$

where L_i is an indicator of whether loan i was liquidated after default, Z_i is a vector of loan characteristics including mortgage note rate, borrower credit score, DTI ratio, and a number of exclusion covariates that help satisfy exclusion restrictions. Our exclusion covariates include combined loan-to-value (CLTV) ratio, 10-year Treasury rate at loan default, a dummy variable for one of the GSEs, and default year dummies. Vintage dummies, state and servicer-fixed effect are also included in the first stage.[18] $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution; γ are unknown parameters to be estimated.

We present the maximum likelihood estimation (MLE) results of this first-stage model in Table 2. Over 2.3 million defaulted loans are used in estimation, among which 898,220 were liquidated.[19] The model coefficients conform to our expectations (i.e., the higher the LTV or the lower the credit score the more likely the loan will end up in default liquidation). The mortgage note rate is shown to be positively associated with the likelihood of default liquidation. Using this model, we are able to correctly classify about 65 percent of the loans in terms of default liquidation versus non-liquidation.

In the second stage, we estimate the following linear regression for loss severity rates by including the inverse Mills ratio we compute based on the first-stage model as a regressor. We also account for potential heteroskedasticity and clustered error terms:

$$y_j = \alpha + X_{j,t}\beta + W_j\eta + \theta\lambda\left(Z_j\hat{\gamma}\right) + \varepsilon_{j,k}, j = 1, \dots, M, \tag{3}$$

[18] The ideal exclusion covariates should have an impact on the selection equation but no direct impact on the loss severity rate. Appendix Table 1 shows the correlation between loss severity rate and selected variables, which helps justify our selection of exclusion covariates.

[19] Here we define default as more than a 90-day delinquency.

where y_j is the loss severity rate of loan i that is liquidated at time t (so only a subset of loans in the first stage). $X_{j,t}$ is a vector of factors that are related to time, including liquidation timelines, the equity position of the property at liquidation, and the macroeconomic environment at the time of liquidation. W_j is a vector of non–time-varying factors, including liquidation type, vintage-fixed effects, state-fixed effects, servicer-fixed effects, and various borrower- and loan-level characteristics. Certain loan characteristics such as combined CLTV ratio, Treasury rate at loan default, Fannie/Freddie dummy and default year dummies are included in the first-stage model but excluded from the second-stage model in order to satisfy the exclusion restrictions. $\lambda\left(Z_j\hat{\gamma}\right)$ is the inverse Mills ratio. Finally, $\varepsilon_{j,k}$ is the heteroskedastic disturbance of loan j clustered by group k. We use the log of default UPB as a weight in the generalized least squares (GLS) estimation given the concern that loss severity rates might be very noisy for small UPB loans. We also assume clustering at the state level, so a two-stage GLS is used in parameter estimation.[20]

Goodman and Zhu (2015) tabulate loss severities of Freddie Mac loans and show loss severity varies with respect to vintage, LTV ratio, credit score, and other dimensions. Our model specification builds on Goodman and Zhu (2015) and other existing literature (see, e.g., Calem and LaCour-Little, 2004; Qi and Yang, 2009). We also incorporate several model enhancements proposed in this paper. Key variables included are contemporaneous (mark-to-market) LTV, liquidation timelines, vintage fixed effects, a distressed sales volume measure (number of default liquidations in a particular metropolitan statistical area (MSA) and a particular quarter normalized by the total number of owner-occupied housing units in that MSA in 2010, plotted in Figure 2), and the timing of liquidation. We include other loan and borrower characteristics we found significant. Loan characteristics include the size of the default UPB in buckets, loan purpose, and property occupancy status.[21] Borrower credit score at origination is also included. We include the type of liquidation; that is, whether the loan is a foreclosure alternative (short sale, deed in lieu of foreclosure, or third-party sales)[22] or a REO sale.

---

[20] We considered alternative county-level clustering but chose the state-level clustering in order not to miss possible error term correlation between properties within the same state but in different counties.

[21] We also test the impact of second liens as the existing literature finds that they present an extra hurdle for foreclosure or foreclosure alternatives (Agarwal et al., 2014b).

[22] Unfortunately, the Freddie data do not break out different types of foreclosure alternatives.

Finally, servicer-fixed effects and state-fixed effects are included to account for potential differences in liquidation efficiency across servicers and different legal environments across states, respectively. For servicer-fixed effects, we have to standardize servicer names and take into consideration M&A activity in the servicing industry, as the mortgage servicing industry has gone through substantial M&A activity in the past 15 years. To do this, we manually collect mortgage servicer M&A data and identify the current servicer of each loan at the time of loan liquidation.

We report the results of our focus variables in Table 3 and include the full model results in Appendix Table 2. As a comparison, we include a baseline specification that does not include the inverse Mills ratio for sample selection correction (model 1). The inverse Mills ratio is highly significant, as shown in model 2, suggesting that there is sample selection bias if we do not include it as an additional regressor in the second-stage regression. We can also see that including the inverse Mills ratio impact other coefficient estimates and helps improve model fit. Our calculation shows that ignoring sample selection will lead to an upward bias of loss severity rates for non-liquidated loans by an average of 0.5 percentage points to 1.0 percentage points.

Mark-to-market LTV (MLTV) is included to measure the equity position of the property at liquidation because it largely determines how much net recoveries investors get from the property sale.[23] It is shown as a significant determinant of loss severity — the higher the MLTV, the higher the severity, consistent with economic intuition and findings in the existing literature (Pennington-Cross, 2003; Calem and LaCour-Little, 2004; Qi and Yang, 2009). There is also non-linearity captured by our MLTV spline function, which is piecewise linear with knots at 100 percent, 120 percent, 140 percent, and 160 percent MLTV shown in Table 3.

Liquidation timelines capture a large portion of the variation in liquidation expenses. As discussed in Section 2, many of the liquidation expenses such as property tax and insurance payments, maintenance and preservation costs and administrative expenses are time related – the longer the timeline the bigger the expense. We document in Section 2 that recent liquidations have significantly longer timelines. Therefore, it is not surprising to see in both models that liquidation

[23] To construct MLTV, we use the county-level house price index from CoreLogic and bring the original LTV current by adjusting the LTV for the change in local HPI. As explained previously, the GSE data only provides three-digit zip codes, so we use a three-digit zip code county cross walk.

timelines are a strong driver of loss severity.[24] A one-year extension of the liquidation timelines could lead to as high as an 8 percentage point (or 20 percent) increase in the loss severity rate, ceteris paribus. Note that the causality here is clear: Lengthening the timeline leads to bigger time-varying liquidation expenses, which are counted as part of the losses as shown in equation (1). In fact, we run a separate regression on liquidation expenses and find liquidation timelines to be the main driver of liquidation expenses.[25]

A potential concern here is that losses are likely higher on loans that have a longer timeline because of unobserved lower quality in the collateral or a slower neighborhood market that a county-level HPI does not control for, rather than the timeline by itself. If that is true, there would be endogeneity bias in the estimation of the effect of liquidation time. To address this concern, we formally test the relation between sales recovery and liquidation timeline. If the endogeneity truly exists, then we would observe a negative relation between sales recovery and liquidation timeline. By running a simple correlation analysis, we find that sales recovery and liquidation timeline only show a weak positive relation. Furthermore, by regressing sales recovery ratio on the liquidation timeline and a few other control variables, we find no significant relation between the liquidation timeline and sales recovery.[26] We thus conclude that the concern of endogeneity is not serious.

We also include dummy variables for loans that were affected by the three major events; namely, the robo-signing scandal, the NMS and the CFPB servicing rules. Later, we will explore how these events affect loss severities through their impact on liquidation timelines, but here we include them to capture any loss severity impact that is not timeline-related (i.e., impact on fixed cost of foreclosure). Loans affected by the robo-signing scandal are defined as those loans serviced by the Big 5 Servicers (Ally, Bank of America, JPMorgan Chase, Citi and Wells Fargo) that fell into delinquency or were foreclosed between September 2010 and October 2012. Loans affected by the NMS are owner-occupied property loans serviced only by the Big 5 Servicers that fell into delinquency or were foreclosed between October 2012 and January 2014. Finally, loans affected

[24] The relation is nonlinear, as evidenced in the significance of the spline functions. As an alternative to spline functions, we experimented with polynomial terms. The results are the same.

[25] For brevity, the liquidation expenses regression results are not reported here. They are available upon request.

[26] For brevity, the correlation and sales recovery regression results are not reported here. They are available upon request.

by the CFPB servicing rules are owner-occupied properties that fell into delinquency or were foreclosed after January 2014. Results in Table 3 suggest that non–timeline-related loss severity rates increased by about 1.0 percent, 0.5 percent and 4.0 percent after the robo-signing scandal, NMS and CFPB, respectively. This is consistent with findings in an industry study that shows the costs of servicing non-performing loans have risen from $482 per loan in 2008 to $1,949 per loan by 2014.[27]

The results of the control variables are consistent with findings in the existing literature or economic intuition. For example, the results in Appendix Table 2 suggest that distressed sales volume, a proxy for local housing market distress, is significant in determining loss severities. This is consistent with the recent literature finding significant discounts for forced sales (see, e.g., Campbell, Giglio, and Pathak, 2011). Importantly, cash-out and rate-and-term refinance loans have higher loss severities, ceteris paribus. This is consistent with findings of appraisal bias, more specifically property value inflation in refinance loans, especially where equity is extracted (Agarwal, Ben-David, and Yao, 2015). Compared with foreclosure alternatives, REO sales result in higher loss severities, even after controlling for liquidation timelines.[28]

In terms of model fit, the adjusted R-squared of our models are about 46 percent, significantly higher than those reported in Lekkas, Guigley, and Van Order (1993) and Calem and LaCour-Little (2004), which are also based on GSE data.[29]

Based on the aforementioned model results, clearly loans liquidated in recent years have benefited from the strong house price recovery in most parts of the country, as more house price appreciation means lower MLTVs. Our model shows that a 10 percentage point decrease in MLTV could be associated with as high as a 6.5 percentage point (or 15 percent) decrease in loss severity rates, ceteris paribus.

---

[27] See Laurie Goodman, "Servicing Costs and the Rise of Squeaky-Clean Loan," *Mortgage Banking*, February 2016.

[28] Foreclosure alternatives in general involve some cooperation from borrowers that serves to lower losses. See Ambrose and Capone (1996) for a discussion of the costs and benefits of foreclosure alternatives.

[29] The R-squared in Lekkas, Guigley, and Van Order (1993) is 6 percent to 7 percent, and that in Calem and LaCour-Little (2004) is 25 percent.

Certainly, as we have shown previously, these recovery benefits are offset by the higher liquidation expenses associated with the extended liquidation timelines. Therefore, to learn the true impact of the shortening/lengthening liquidation timelines, we run simulations based on the loss severity model estimated. In our simulation, for each loan, we shift the liquidation date by one year. We then recalculate other time-related variables such as the MLTV, distressed sales volume, and the robo-signing scandal, NMS, and CFPB dummies. Finally, we score each loan with our loss severity model and the simulated regressors. We then calculate the change in loss severity rates under the scenarios of a one-year reduction and a one-year extension of liquidation timelines.

Results aggregated at the state level are shown in Table 4. Since different states have experienced different house price appreciation trajectories, the impact of the shortening or lengthening of liquidation timelines has different impacts across states. For example, Florida had a strong house price recovery in recent years; therefore, shortening the liquidation timeline provides a smaller benefit than states with less appreciation. In terms of further lengthening liquidation timelines, the damage is relatively small to California loans; as they would have the chance to benefit from more house price appreciation if liquidation is delayed further into the future. We compare our simulated loss severities with the baseline results in Figure 4 for all states. Again, we see different impacts in different states. Cutting foreclosure timelines by one year could result in a benefit that ranges from 5 percentage points to 8 percentage points, depending on the tradeoff between lower liquidation expenses and less sales recoveries.

4. Factors Contributing to Liquidation Timeline Extensions

We have established that extended liquidation timelines contribute to higher loss severities post-crisis. Now we explore what drives the lengthy timelines.

First, we break out the liquidation timeline into REO timelines and foreclosure timelines. As described in Appendix Figure 1, before a default property becomes an REO, it has to go through the foreclosure process. The length of the foreclosure process depends on factors such as the court process in judicial states and loan modification requirements by regulators. Once a property becomes REO, the lender/investor has full control, except for redemption requirements in some statutory redemption states.

Figure 5 shows the average foreclosure timelines and REO timelines. REO timelines are quite stable. The time series variations of liquidation timelines come from changes in foreclosure timelines. Recent research claims that, among various factors, enactment of the new mortgage servicing rules by NMS and CFPB might have played an important role in extending foreclosure timelines (see, e.g., Cordell et al., 2015; Cordell and Lambie-Hanson, 2016). We conduct multivariate analysis and difference-in-differences tests to explore these factors.

4.1. Accelerated Failure Time Model for Liquidation Timelines

We first estimate a set of liquidation timeline models that are similar to that in Cordell et al. (2015). The model is an accelerated failure time model that assumes a log normal distribution of the liquidation timeline.[30] A key issue addressed with this model is the censoring problem: For loans that recently went into delinquency but have not been liquidated at our data cutoff point we simply do not observe the true timeline (i.e., the timeline is right-censored). Our sample contains all loans that have been 90 or more days delinquent,[31] including those for which we observe the true timeline because of liquidation (uncensored) and those we do not observe the true timeline (censored).[32] The model takes the following form:

$$\log(\mathrm{T_i}) = \mu + \mathrm{U_i}\delta + \sigma \cdot \varepsilon_\mathrm{i}, \quad (4)$$

where T_i is the observed timeline for uncensored loans or duration for censored loans, U_i is a vector of covariates, and ε_i is a random value from standard normal distribution. μ, δ and σ are unknown parameters. A maximum likelihood estimation method with a ridge-stabilized Newton-Raphson algorithm is used in the model estimation (see Kalbfleisch and Prentice, 1980).

In Figure 6, we plot the average liquidation timelines together with the share of censored loans. Unsurprisingly, the share of censored loans increases significantly as we get closer to the data

[30] Alternative distributional assumptions were tested in Cordell et al. (2015), and they found these other distributions produced similar results.

[31] For loans that have been in and out of 90-day delinquency more than once, we only include the last 90-day delinquency in our sample to avoid the complication arising from re-default or modification.

[32] Loans paid off or cured after they became 90-day delinquent are included in our sample as censored observations. For those loans, we treat the date of the payoff or the date of returning to the "current" status as the censoring date.

cutoff date, which causes the average liquidation timeline to decline significantly as we approach the data cutoff point. Our failure time model is able to capture the censoring effect.

In terms of model covariates, we include the usual loan and borrower characteristics, such as the mark-to-market LTV at the time of serious delinquency, local HPI trends, state legal environments, and dummies for default timing. An innovation we have in this paper is the inclusion of a state-level foreclosure pipeline measure that is constructed as the total number of loans that are 60 or more days delinquent or in the foreclosure process, divided by the total number of active loans in a particular state and quarter. This foreclosure pipeline measure is meant to capture the effect of court and servicing capacity constraints and congestion.

We present our main timeline model results in Table 5 and report the full model results in Appendix Table 3. Given our purpose, we focus our discussion on the default timing effects. As we see in Table 5, we divide our whole study period into six subperiods. In our model, the pre-crisis, pre-February 2007 is the reference period, the period from February 2007 to October 2008 marks the first part of the housing crisis, and November 2008 to August 2010 marks the period of a GSE foreclosure moratorium as well as the pre-robo-signing period when the Obama administration's foreclosure-prevention initiatives were tested. The final three subperiods are marked by the following events: September 2010 is when the robo-signing scandal broke out, freezing foreclosures at the largest servicers (and some others); October 2012 is the deadline for the Big 5 Servicers to incorporate the NMS servicing standards; and January 2014 is when the new CFPB servicing rules took effect.

We see from Table 5 that, after controlling for housing market conditions, the foreclosure pipeline, and other loan and borrower characteristics, liquidation timelines have increased over time, especially in the post-crisis periods. Note that the coefficients on the timing dummies are largest for loans that became seriously delinquent in the two most recent periods, periods driven by recent legal and regulatory changes, namely the NMS and new CFPB servicing rules. We now turn our attention to estimating the effects of these changes.

4.2. Difference-in-Differences (DID) Tests

There are several reasons why we see prolonged liquidation timelines post-crisis, especially in judicial states. First, there are the legacy loans coming principally from various moratoria and the

Robo-signing scandal in 2010 that resulted in the freezing of foreclosures. We observe from the Freddie Mac data that delinquent loans subject to the GSE moratoria starting in November 2008 on average have a liquidation timeline five months longer than those liquidated before the GSE moratoria.[33]

Second, the loan servicing industry has gone through substantial changes post-crisis, especially after the breakout of the robo-signing scandal in September 2010. Cordell et al. (2015) and Cordell and Lambie-Hanson (2016) document some regime shifts in the servicing industry and the resulting prolonged liquidation timelines.[34] Two important national legal remedies stand out. First, the NMS announced in March 2012 that it not only required the five largest mortgage servicers[35] in the country to pay out cash compensations to some of the borrowers they serviced but also required them to comply with over 300 new servicing rules, especially in their handling of delinquent loans.[36] In January 2013, the CFPB announced new servicing rules to take effect a year later in January 2014.

Many of these changes designed to improve loan-modification opportunities for borrowers have extended foreclosure timelines or otherwise increased LGDs. As Cordell and Lambie-Hanson (2015) show, the CFPB rules prohibit servicers from beginning foreclosure proceedings on owner occupants until they are more than 120 days delinquent. Prior to 2014, over half of all loans were referred to foreclosure before 120 days; after 2014, only 2 percent were.

Prior to the NMS rules, servicers engaged in foreclosure alternatives while the foreclosure process was proceeding, a practice known as *dual tracking*. This resulted in many borrowers being denied loan modifications, especially in non-judicial states with fast timelines. The new rules specifically prohibit this practice, requiring servicers to commence with foreclosure proceedings only *after* modification options have been exhausted, lengthening timelines in general but by large amounts

[33] These loans include those that were delinquent in November 2008 when the GSE moratoria commenced and all loans delinquent after 2008 up to the point of the robo-signing scandal, at which time an entirely new event occurred, resulting in the outright freezing of foreclosures at the largest servicers.

[34] These papers use the McDash data that cover most loans serviced by the top 10 mortgage servicers from 2005 to 2014.

[35] The five largest mortgage servicers are Ally, Bank of America, JPMorgan Chase, Citi, and Wells Fargo.

[36] Later there were a series of additional settlements, such as the Ocwen Settlement (December 2013) and the National Suntrust Settlement (January 2016).

in select contested cases. Finally, because so many borrowers during the crisis found it difficult to communicate with servicers during the long workout process, the new rules require servicers to designate a "single point of contact" for each borrower after they become delinquent, increasing overhead costs at servicers.[37] Those changes resulted in more loan modifications but also in much higher servicing costs[38] by 2015 and, as we will see, increased liquidation timelines.

One issue with the two aforementioned studies is that they base their claims of regime shifts on time dummies and inspection of the data, not on statistical inference. To test empirically their claim of regime shifts in foreclosure timelines in our context, we conduct several formal tests.

We conduct our tests in a difference-in-differences (DID) framework so we can more precisely target loans treated under the new servicing rules. Our DID tests are in the standard form:

$$y = \alpha + \beta_1 T + \beta_2 P + \beta_3 T \cdot P + W\gamma + \varepsilon , \quad (5)$$

where y is foreclosure timelines (in log); T represents the treatment group, which are loans targeted by a new policy or affected by an event; P represents post-treatment, which is the period after policy implementation or event occurrence; W is the vector of control variables; and ε is the error term. Here β_1 captures the generic difference between the treatment group and the control group that is not related to time; β_2 captures the general time trend that applies to both the treatment group and the control group; and β_3 captures the treatment effect, which can be interpreted as the impact of the policy or event.

The two events we test are the NMS and the new CFPB servicing rules, whose rules comprise the new regime for mortgage servicing. In the NMS, a comprehensive set of servicing rules (more than 300 rules) were introduced, which we hypothesize lengthened foreclosure timelines. For example, the NMS required that servicers must evaluate borrowers for all available loan modification options before referring borrowers to foreclosure. It also required that servicers may not proceed

---

[37] Consumer Financial Protection Bureau. 2013. Summary of Final Mortgage Servicing Rules, available at http://files.consumerfinance.gov/f/201301-cfpb servicing-rules summary.pdf .

[38] According to the Mortgage Bankers Association, while servicer compensation did not change between 2008 and 2015, the cost to service performing loans tripled (from \$59 to \$181) while the cost to service non-performing loans increased fivefold (from \$484 to \$2,386). See Urban Institute, 2016, "The Mortgage Servicing Collaborative" found at https://www.urban.org/sites/default/files/msc factsheet .pdf.

with foreclosure sales while an appeal of the denial of a loan modification is pending.[39] The CFPB servicing rules incorporated the NMS rules and added additional rules. As mentioned, under the CFPB rules, servicers cannot start the foreclosure process on owner-occupied property loans until the 120th day of delinquency, replacing the old standard of 75 to 90 days (sooner for properties found vacated).[40, 41]

Only owner-occupied property loans, not investment properties, are the target of the NMS and CFPB rules, which provides an opportunity to establish a treatment group and a control group. In addition, the NMS rules applied principally to the Big 5 Servicers under the settlement, while the CFPB rules generally apply to all servicers, only exempting very small servicers from certain requirements.[42] The NMS was announced in March 2012, with a date of October 2, 2012, to be in full compliance with its 304 servicing standards. The CFPB's servicing rules were announced in January 2013, with an effective date of January 2014. Thus, we select October 2012 and January 2014 as the event dates for our NMS and CFPB tests, respectively.[43]

An important assumption of the DID approach is that the treated and the control groups follow a parallel trend prior to the event. To test the validity of this parallel trend assumption, we interact the treatment variables with time dummies in regressions to show how the trends progress period-by period before and after the event. The charts in Appendix Figure 3 plot the coefficients four quarters before and four quarters after the CFPB policy change (event date). Results clearly show

---

[39] The NMS also limited fees that a servicer can collect from a defaulted borrower.

[40] On the cost side, the CFPB rules require a servicing employee be assigned no later than the 45th day of delinquency, with that employee becoming a "single point of contact" for the borrower while the loan is delinquent, tightening the initial rules in the NMS. Previously, no such rule existed.

[41] See National Housing Resource Center, "Understanding the National Mortgage Settlement," June 2013 and Bureau of CFPB, "Amendments to the 2013 Mortgage Rules under the Real Estate Settlement Procedures Act (Regulation X) and the Truth in Lending Act (Regulation Z)," 2016.

[42] Small servicers are defined as those that service 5,000 or fewer loans that they or an affiliate service.

[43] Note that, for our study purposes, the effective date of the policy change is more appropriate to use than the announcement date. This is because servicers, not like stock market participants who respond to news, were bounded by the effective date but not the announcement date – they can wait until the effective date to make changes. One could argue that some servicers started to make a change as soon as they heard the announcement. However, that type of action will cause a smaller difference we could observe in the data between the pre- and post-event periods, and work against us finding significant results.

that our assumption of parallel trend prior to the event is valid and that there is a clear structural break showing the impact of the policy change.

Table 6 presents our DID test results on foreclosure timelines. In Panel A, we use owner-occupied property loans as the treatment group and investment properties as the control group. Because the robo-signing scandal and subsequent NMS were mainly the result of the Big 5 Servicers failing to file proper documentation with courts, it applied principally to defaulted loans of the Big 5 Servicers in judicial states, so we focus these tests on judicial states. To avoid confounding effects of other events, we limit our test window. For the NMS test, the test window extends from September 2010, which is the start of the robo-signing scandal, to January 2014, which is the effective date of the new CFPB servicing rules. Recall that the NMS event date is October 2012. Therefore, we limit the pre-event loans, which are our control group, to those that were seriously delinquent and foreclosed between September 2010 and October 2012, and post-event loans are those that were seriously delinquent and foreclosed between October 2012 and January 2014. Furthermore, we eliminate from our sample loans that fell into serious delinquency before October 2012 but were liquidated after October 2012 because they were exposed to both the pre- and post-event environment. Therefore, in Panel A, our treatment group affected by the NMS rules is owner-occupied property loans in judicial states that were serviced by the Big 5 Servicers and exposed to the post-NMS environment. For the CFPB test, we limit our test window from October 2012 to September 2015 with the event date of January 2014.

In Panel B, we use delinquent loans serviced by the Big 5 Servicers as the treatment group and those serviced by smaller servicers as the control group. Following the same logic of Panel A, to avoid confounding effects, we now limit our tests to owner-occupied property loans in judicial states.

The results in the first column of Table 6, Panel A suggest that although there is a general trend of longer foreclosure timelines post-NMS (the significant and positive coefficient of *Post-event* of 0.278), the treatment group has a larger increase in foreclosure timelines post-NMS than the control group (the significant and positive coefficient of *Owner* × *Post-event* of 0.264). This DID result provides evidence that the NMS had a causal impact on foreclosure timelines. Similarly, we see from the second column that there is a general trend of longer foreclosure timelines post-CFPB because the treatment group was more affected (the significant 0.231 coefficient).

The results in the first column of Panel B tell a similar story: Post-NMS, the treatment group, loans serviced by the Big 5 Servicers, had higher increases in foreclosure timelines than the control group, loans serviced by smaller servicers. This is shown by the significant and positive coefficient of 0.136 on our treatment group of owner-occupied property loans at the Big 5 Servicers. This is again evidence of a causal impact of NMS on foreclosure timelines.

The results in the second column of Table 6, Panel B are also of interest. We see again that there is a general trend of longer foreclosure timelines post-CFPB (significant and positive coefficient of *Post-event* of 0.625). Loans serviced by the Big 5 Servicers, which are the treatment group in the NMS period, have shorter changes in timelines post-CFPB than loans serviced by smaller servicers (the significant and negative coefficient of *Big 5 Servicer × Post-event* of -0.118).[44] Note that in this case the smaller servicers effectively *become* the treatment group because they are the ones now most affected by the new servicing rules. This is also consistent with the hypothesis that the new CFPB servicing rules had a causal impact on foreclosure timelines beyond the general trend captured by post-event coefficient of 0.625. The logic is as follows: If the implementation of the CFPB servicing rules represents a shock to servicers, it is more of a shock to smaller servicers other than the Big 5 Servicers, which had already implemented the NMS rules in 2012. As noted, because the CFPB rules built on the NMS rules, we hypothesize that the CFPB servicing rules have limited incremental impact on the Big 5 Servicers. In contrast, the rules should have a much larger impact on the smaller servicers because they were not required to adopt the new rules before January 2014. This is exactly what we see in the second column of Table 6 Panel B with the negative coefficient on the Big 5 Servicers.[45]

We also use the GSE foreclosure moratoria to conduct a placebo test. The thought process is that GSE foreclosure moratoria applied to all loans, no matter if they are serviced by the Big 5 Servicers or smaller servicers and no matter if they are owner-occupied property loans or investment loans.

[44] Two non-NMS servicers, Ocwen (March 2014) and Suntrust (September 2014), engaged in separate settlements and thus are excluded from this sample.

[45] There is potentially a selection effect in terms of whether loans that default cure or prepay with or without modification. It may be that there have been changes over time in what kinds of loans get modified or cured which are driving the change in trends in liquidation time between the treatment and control group. The placebo test discussed next helps address this concern.

Therefore, we should not expect a treatment (DID) effect. Results in Table 7 confirm that this is exactly what we see in the data.

In summary, the DID tests suggest causal impacts on loss severities from the NMS and CFPB servicing rules extending foreclosure timelines. Note that the DID results presented in the aforementioned tables should be treated as *lower bound estimates* of the impact of those events. This is because there could be spillover effects between the treatment group and the control group, and certain servicers could have adopted the rules in anticipation of their taking effect.[46] Thus, the tests are biased *against* finding significant differences because the effects caused by the freezing of foreclosures between September 2010 and October 2012 that impacted our control group also raised timelines for them.

5. Conclusions and Discussion

Despite its importance, research work on mortgage LGD has been sparse until now, mainly because of data limitations. Thus, the release of GSE detailed loan-level loss data provides a unique opportunity to study residential mortgage severities, especially because GSE data contain detailed information on the components of loss severity encompassing the full boom, bust, and recovery periods during a period of profound change in U.S. mortgage markets. Our analysis generates a number of findings that have important implications for our understanding of the financial crisis and the future of the housing finance system in the U.S.

First, the rise of loss severities from the early 2000s to the crisis period is substantial. Loss severity rates of defaulted and liquidated GSE fixed-rate mortgages more than doubled; we expect that increases in severities on their nontraditional mortgages not included in this sample were even higher. Although research on the downfall of the GSEs has focused on the sharp rise in mortgage defaults, no study has systematically studied the contribution of loss severities tied to these defaults. A doubling of loss severities on their traditional business clearly played a role, so it will be important to make this a part of future research on the insolvency of the GSEs during the financial crisis.

[46] Please also refer to footnote 44.

Second, our research shows that persistently high loss severities post-crisis is due to a combination of factors, including legacy loan overhang, prolonged foreclosure timelines, and increased foreclosure expenses, which all offset higher sales recoveries as house prices rebounded. In our DID tests, we focus on more recent defaults to show how the new servicing standards stemming from the NMS and CFPB servicing rules have contributed to longer timelines and higher LGDs. Under the new rules, liquidation expenses are higher because of policy changes that resulted in longer foreclosure timelines that increased the cost of servicing delinquent mortgages.

More important, policy analysts now claim higher servicing costs are affecting the provisioning of mortgage credit on the origination side, giving rise to the "squeaky clean" loan with minimal risk of ever going delinquent, potentially disadvantaging borrowers with limited or weaker credit.[47] Our analysis supports the hypothesis that the new servicing rules are contributing to higher servicing costs. However, these analyses only focus on costs; the new servicing rules might have benefited borrowers by providing more opportunities for loan modifications, allowing more borrowers to stay in their homes (Cordell and Lambie-Hanson, 2016). Therefore, more analysis is needed to examine the cost–benefit tradeoffs of the new CFPB servicing rules.

On a practical level, all of these factors will play into the pricing of credit risk for the GSEs' fast-growing credit risk transfer (CRT) business. To that end, models developed in this paper can help the investment community more accurately estimate losses. Our models can also be used in stress testing and other applications.

Finally, our findings imply that losses can be underestimated by not considering changes taking place in the U.S. mortgage market. From this perspective, our results support the notion that statistical models can fail if they do not incorporate specific institutional settings and structural breaks in models, as argued by An et al. (2012); Rajan, Seru, and Vig (2015); and An, Deng and Gabriel (2019).

[47] See Laurie Goodman, "Servicing Costs and the Rise of the Squeaky-Clean Loan," Mortgage Banking, February 2016.

References

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff. 2014a. "Predatory Lending and the Subprime Crisis." *Journal of Financial Economics*, 113(1): 29–52.

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Yan Zhang. 2014b. "Second Liens and the Holdup Problem in Mortgage Renegotiation." NBER Working Paper No. 20015.

Agarwal, Sumit, Itzhak Ben-David, and Vincent Yao. 2015. "Collateral Valuation and Borrower Financial Constraints: Evidence from the Residential Real Estate Market." *Management Science*, 61(9): 2220–2240.

Ambrose, Brent W., and Charles A. Capone, Jr. 1996. "Cost-Benefit Analysis of Single Family Mortgage Foreclosure Alternatives." *Journal of Real Estate Finance and Economics*, 13(2): 105–120.

An, Xudong, Yongheng Deng, Eric Rosenblatt, and Vincent W. Yao. 2012. "Model Stability and the Subprime Mortgage Crisis." *Journal of Real Estate Finance and Economics,* 45(3): 545–568.

An, Xudong, Yongheng Deng, and Stuart A. Gabriel. 2019. "Default Option Exercise over the Financial Crisis and Beyond." Forthcoming in *Review of Finance*.

Berkovec, James A., Glenn B. Canner, Stuart A. Gabriel, and Timothy H. Hannan. 1998. "Discrimination, Competition, and Loan Performance in FHA Mortgage Lending." *Review of Economics and Statistics,* 80(2): 241–250.

Calem, Paul S., and Michael LaCour-Little. 2004. "Risk-based Capital Requirements for Mortgage Loans." *Journal of Banking and Finance,* 28, 647–672.

Campbell, John Y., Stefano Giglio, and Parag Pathak. 2011. "Forced Sales and House Prices." American Economic Review, 101 (5): 2108-2131.

Campbell, Tim S. and J. Kimball Dietrich. 1983. "The Determinants of Default on Conventional Residential Mortgages." *Journal of Finance* 48(5): 1569-1581.

Capozza, Dennis, and Thomas Thomson. 2005. "Optimal Stopping and Losses on Subprime Mortgages." *Journal of Real Estate Finance and Economics,* 30(2): 115–131.

Clapp, John M., Yongheng Deng and Xudong An. 2006. "Unobserved Heterogeneity in Models of Competing Mortgage Termination Risks." *Real Estate Economics*, 34(2): 243-273.

Cordell, Larry, and Lauren Lambie-Hanson. 2016. "A Cost-Benefit Analysis of Judicial Foreclosure Delay and a Preliminary Look at New Mortgage Servicing Rules." *Journal of Economics and Business,* 84: 30–49.

Cordell, Larry, Liang Geng, Laurie Goodman, and Lidan Yang. 2015. "The Cost of Foreclosure Delay." *Real Estate Economics,* 43(4): 916–956.

Demyanyk, Yulia, and Otto Van Hemert. 2011. "Understanding the Subprime Mortgage Crisis." *Review of Financial Studies*, 24(6): 1848–1880.

deRitis, Cristian and Deniz K. Tudor. 2016. "CECL's Implications for Bank Profitability, System Stability, and Economic Growth." *Moody's Analytics Risk Perspective*.

Elul, Ronel, Nicholas S. Souleles, Souphala Chomsisengphet, Dennis Glennon, and Robert Hunt. 2010. "What 'Triggers' Mortgage Default?" *American Economic Review,* 100(2): 490–494.

Goodman, Laurie, and Jun Zhu. 2015. "Loss Severity on Residential Mortgages: Evidence from Freddie Mac's Newest Data." *Journal of Fixed Income* 25(2): 48-57.

Kalbfleisch, John D., and Ross L. Prentice. 1980. *The Statistical Analysis of Failure Time Data*. New York: John Wiley & Sons.

Keys, Benjamin, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig. 2010. "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans." *Quarterly Journal of Economics*, 125(1): 307–362.

Layton, Don. 2020. "Demystifying GSE Credit Risk Transfer. Part 1-What Problem Are We Trying to Solve?" Joint Center for Housing Studies, Harvard University.

Le, Ben and Anthony Pennington-Cross. 2018. "Mortgage Losses: Loss on Sale and Holding Costs." 2019 AREUEA Annual Conference Paper.

Lekkas, Vassilis, John M. Quigley, and Robert Van Order. 1993. "Loan Loss Severity and Optimal Mortgage Default." *Journal of the American Real Estate and Urban Economics Association,* 21 (4): 353–371.

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

Pennington-Cross, Anthony. 2003. "Subprime & Prime Mortgages: Loss Distributions." OFHEO Working Paper 03-1.

Qi, Min, and Xialong Yang. 2009. "Loss Given Default of High Loan-to-Value Residential Mortgages." *Journal of Banking and Finance,* 33: 788–799.

Rajan, Uday, Amit Seru, and Vikrant Vig. 2015. "The Failure of Models that Predict Failure: Distance, Incentives and Defaults." *Journal of Financial Economics,* 115(2): 237–260.

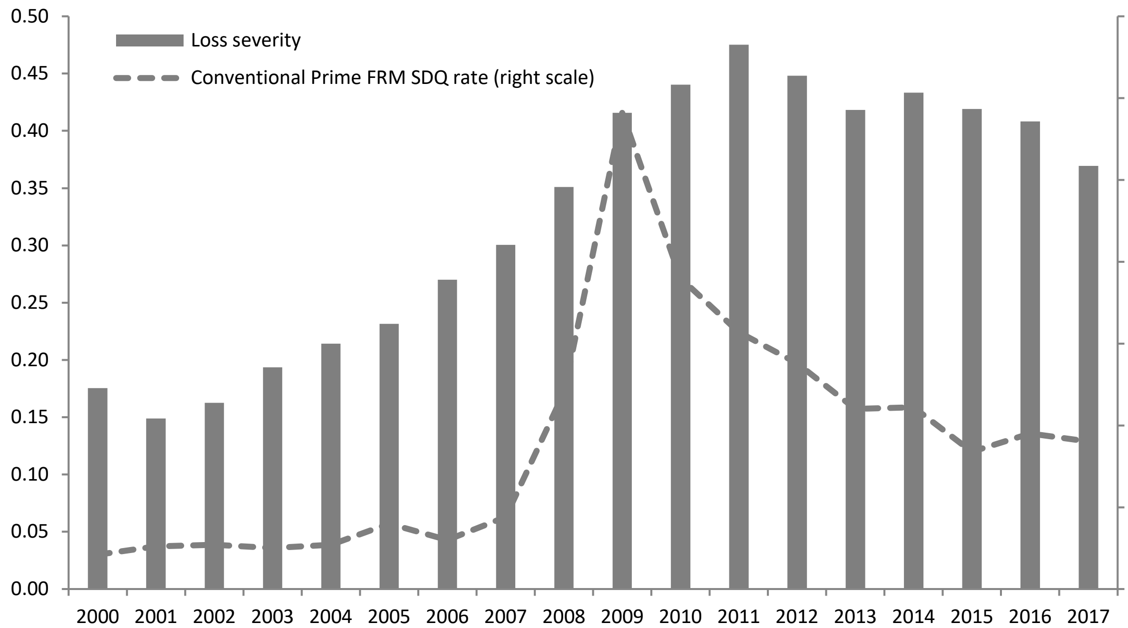


Figure 1: Loss Severities of GSE Loans by Year of Liquidation

Notes: The bars depict weighted average loss severity rates of loans liquidated in the calendar year (left scale); the dashed line represents serious delinquency (SDQ) rates of conventional prime fixed-rate mortgage loans (right scale). Loss severity numbers are calculated based on the Freddie Mac Single-Family Loan-Level Dataset and Fannie Mae Single-Family Loan Performance Data (hereafter, Freddie Mac data, and Fannie Mae data, respectively , and government-sponsored enterprise (GSE) data altogether), and the SDQ data are from Mortgage Bankers Association's National Delinquency Survey through Haver Analytics. GSE loans included here are first-lien, full-documentation, and fully amortizing fixed-rate mortgages for single-family homes that were liquidated through short sale, foreclosure sale or real-estate owned (REO) sale. Repurchase loans are excluded. These loans were originated between 1999 and 2016 and liquidated between 2000 and 2017. Loss severity is defined as (net sales proceeds + non-MI recoveries – UPB –expenses) ÷ UPB. Expenses include all allowable expenses that the GSEs bear in the process of acquiring, maintaining, and/or disposing of properties. Non-MI recoveries include tax and insurance refunds, hazard insurance proceeds, rental receipts, and positive escrow. MI recoveries or carrying costs are not considered in our calculation. UPB = unpaid principal balance; MI = mortgage insurance.

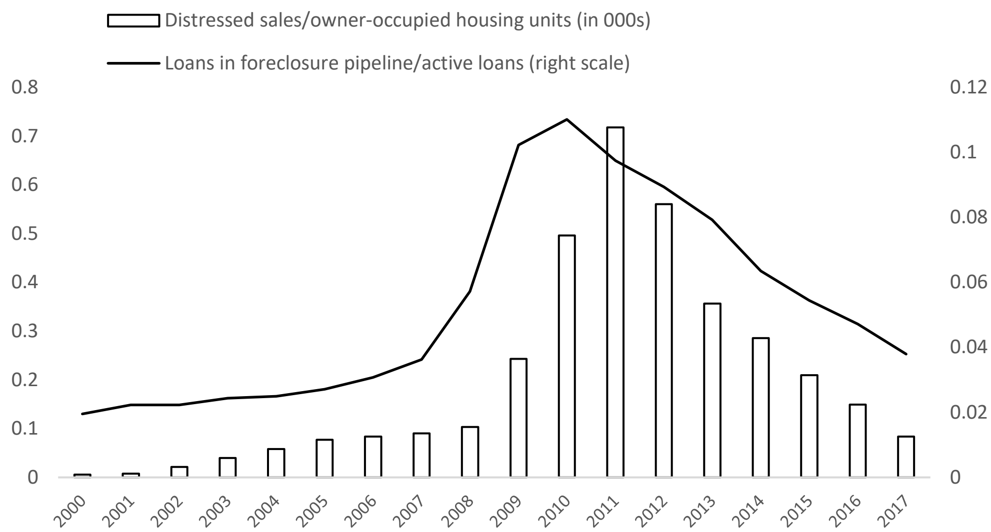


Figure 2: Liquidation Volume and Foreclosure Pipeline

Notes: The bars depict the average number of distressed sales normalized by the total number of owner-occupied housing units (in 000s) across all metropolitan statistical areas (MSAs) in each year. The line shows the average number of loans 60 or more days delinquent or in foreclosure normalized by the total number of active loans across all states in each year. The former series is calculated based on the GSE data and census data, while the latter is calculated based on McDash data.

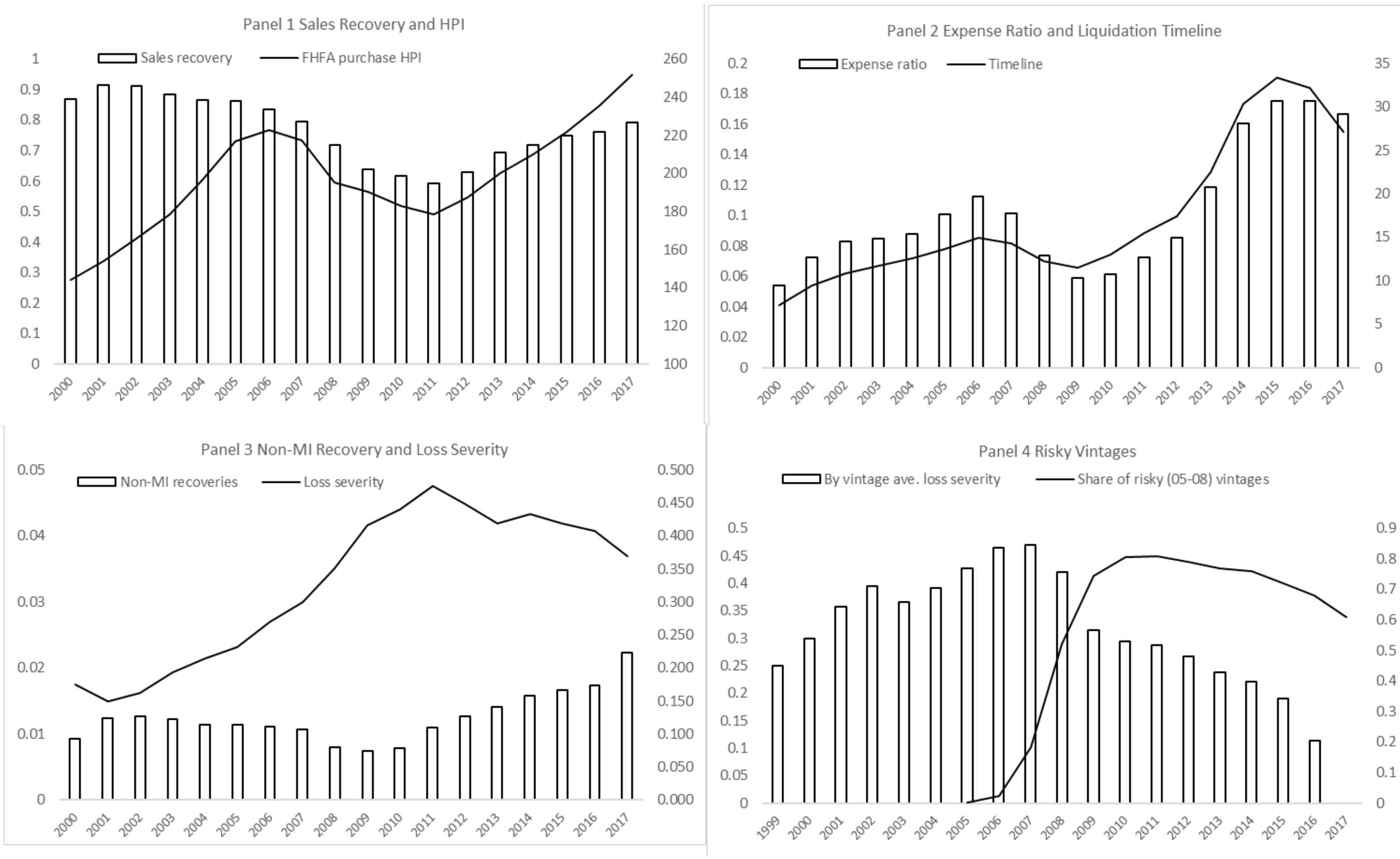


Figure 3: Sales Recovery, Expenses, Risky Vintages, and Liquidation Volume

Notes: Based on the GSE data. For loans liquidated through short sales, foreclosure sales or real-estate owned (REO) sales. Panel 1 shows the average sales recoveries as a percentage of default unpaid principal balance (UPB) and the FHFA home-purchase house price index (HPI); Panel 2 shows the average liquidation expenses as a percentage of default UPB and the average liquidation timeline; Panel 3 shows the average non-MI recoveries as a percentage of default UPB and the average loss severity rates; and Panel 4 shows the raw loss severity by vintage as well as the percentage of liquidations from the risky (05-08) vintages.

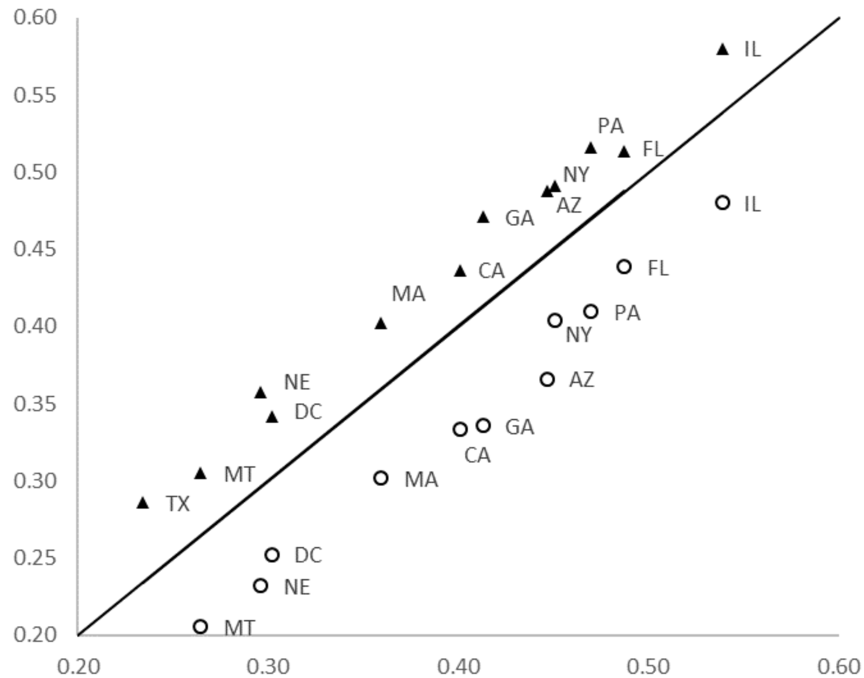


Figure 4: Loss Severity Simulation Results for Selected States

Notes: Results are based on the regression dataset. The triangles show the simulated versus model-fitted default balance weighted average loss severities. The simulation is based on our loss severity model and assumes that for each loan the liquidation timeline is one-year longer than the actual timeline. The circles show similar results except that the simulation assumes a period one-year shorter than actual liquidation timelines. We only show selected states in the graph. For full results, see Table 4.

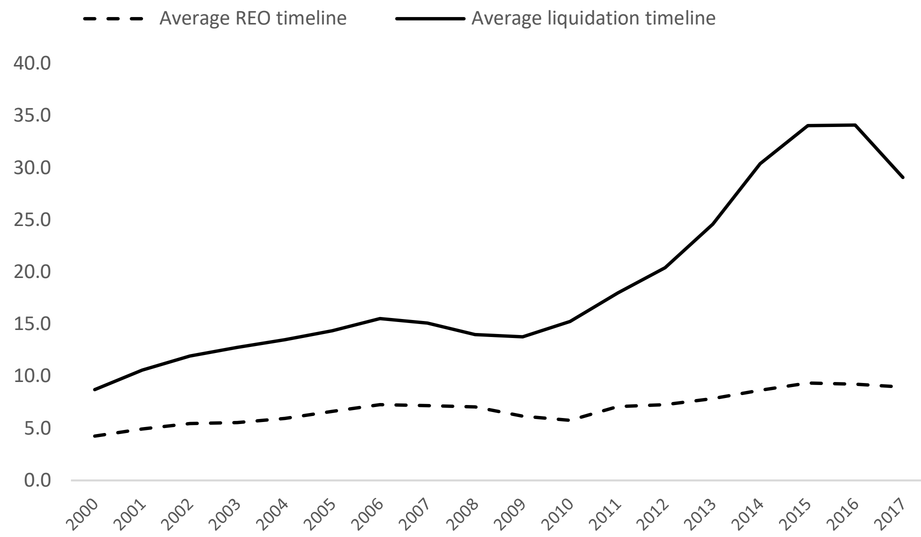


Figure 5: Average Foreclosure and REO Timelines for Liquidated Loans

Notes: Based on the GSE data. The solid line shows the average liquidation timeline, and the dashed line shows the in- real-estate owned (REO) timeline, therefore, the difference between the two lines are foreclosure timelines. The liquidation timeline is calculated as the number of months from serious delinquency (SDQ) to final liquidation, either in the form of REO sales or foreclosure alternatives such as short sales.

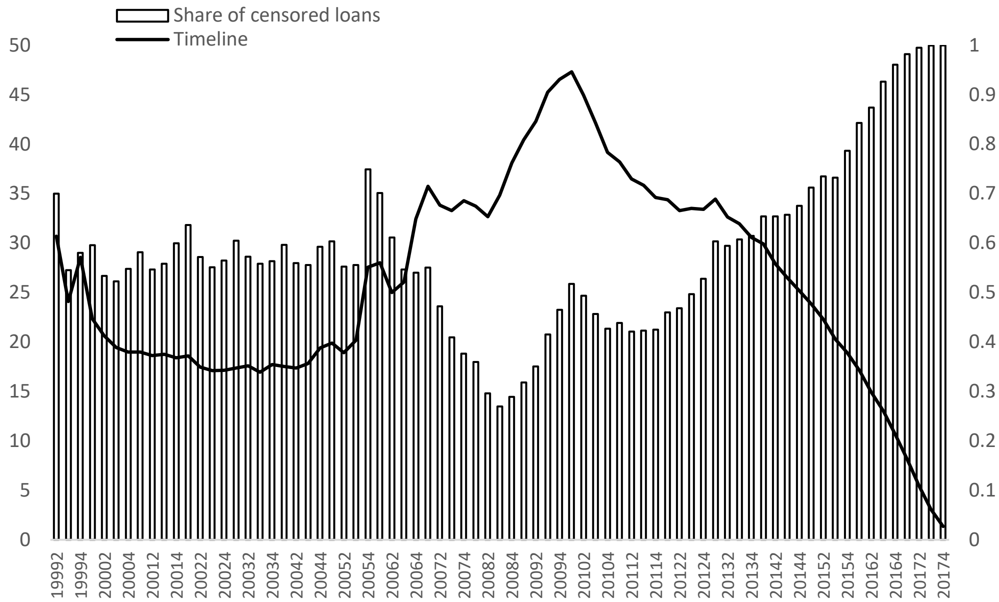


Figure 6: Average Liquidation Timeline and Share of Censoring Among Serious Delinquent Loans

Notes: Based on the GSE data. X-axis is serious (90-day) delinquency date. The dark line shows the default balance weighted average liquidation timelines, while the bars show the share of serious delinquent loans that were not liquidated through short sale, foreclosure sale or real-estate owned (REO) sale (could be prepaid, unresolved or REO not liquidated). Calculation is based on all serious delinquent loans.

Table 1: Summary Statistics

Variable	Liquidated Loans Mean	Liquidated Loans Median	Liquidated Loans Std. Dev.	All Loans Mean	All Loans Median	All Loans Std. Dev.
Original loan amount	166,894	149,000	90,915	193,045	170,000	106,530
Mortgage interest rate	6.33	6.25	0.73	5.35	5.38	1.33
Loan-to-value (LTV) ratio	80.4	80.0	10.3	74.8	78.0	12.2
Borrower credit score	692	690	55	739	750	53
Debt-to-income ratio	39	40	12	34	34	11
Home purchase loan	0.35	0.00	0.48	0.39	0.00	0.49
Cash-out refinance loan	0.39	0.00	0.49	0.27	0.00	0.45
Rate/term refinance loan	0.25	0.00	0.44	0.33	0.00	0.47
Owner-occupied	0.88	1.00	0.33	0.90	1.00	0.31
Second home	0.04	0.00	0.19	0.04	0.00	0.20
Investment property	0.09	0.00	0.28	0.06	0.00	0.24
Loan in judicial state	0.42	0.00	0.49	0.39	0.00	0.49
Default unpaid balance	157,778	139,706	88,782	–	–	–
Mark-to-market LTV (MLTV)	93.4	89.1	28.4	–	–	–
Liquidation timeline (months)	19	16	14	–	–	–
Current quarter distressed sales volume	0.38	0.38	0.33	–	–	–
REO liquidation (vs. other foreclosure alternatives)	0.69	1.00	0.46	–	–	–
Loss severity	0.42	0.41	0.29	–	–	–
Number of loans		898,220			52,122,741	

Notes: Based on the government-sponsored enterprise (GSE), CoreLogic, and census data. Loans included here are first-lien, full-documentation, and fully amortizing fixed-rate mortgages for single-family homes that were liquidated through short sales, foreclosure sales or real-estate owned (REO) sales. Repurchase loans are excluded. These loans were originated between 1999 and 2016 and liquidated between 2000 and 2017. Bad data and outliers (e.g., those with extremely high loss severity and those with missing loan-to-value (LTV) ratio information) are excluded. Mark-to-market LTV is calculated based on original loan amount and change in house price index (HPI) between loan origination and liquidation. CoreLogic county-level HPI is used (the three-digit zip code contained in the GSE data is mapped to county). The liquidation timeline is the number of months lapsed between serious (90-day) delinquency and loan liquidation. Distressed sales volume is the number of default liquidations in a specific metropolitan statistical area (MSA) normalized by the total number of owner-occupied housing units (in 000s) in that MSA. Loss severity is defined as (net sales proceeds + non-MI recoveries – UPB – expenses) ÷ UPB. Expenses include all allowable expenses that GSEs bear in the process of acquiring, maintaining, and/or disposing of properties, even though property selling expenses such as commission is taken into consideration separately (in net sales proceeds). Non-MI recoveries include tax and insurance refunds, hazard insurance proceeds, rental receipts, and positive escrow. MI recoveries or carrying costs are not considered in our calculation. Loss severity mean is weighted by default balance. UPB = unpaid principal balance; MI = mortgage insurance.

Table 2: MLE Estimates of the First Stage of the Heckman Two-Stage Loss Severity Model

Covariate	Estimate (S.E.)
Combined loan-to-value (CLTV) ratio at origination	0.030***
	(0.001)
GSE A	-0.188***
	(0.002)
10-year Treasury rate at default	-0.067***
	(0.002)
Default year-fixed effect	Y
Loan and borrower characteristics	Y
Vintage-fixed effect	Y
State servicer-fixed effect	Y
N	2,334,780
-2LogL	2,647,504
AIC	2,647,738
Percent concordant	74.3

Notes: Based on the GSE data. These are maximum likelihood estimation (MLE) estimates from a Probit model (equation 2) where the dependent variable is an indicator of whether the loan was liquidated through short sale, foreclosure sale or real-estate owned (REO) sale. Loan and borrower characteristics included are mortgage coupon rate, borrower credit score, debt-to-income (DTI) ratio, property occupancy type, and loan purpose. Standard errors are in parentheses. All defaulted (serious delinquent) loans, both liquidated and non-liquidated, are included except for those with bad data. GSE A refers to one of the GSEs. *** for p<0.01%, ** for p<0.1%, and * for p<5%.

Table 3: GLS Estimates of the Second Stage of the Heckman Two-Stage Loss Severity Model

Covariate		Estimate(S.E.)	
		Model 1	Model 2
Inverse Mills ratio			-0.063***
			(0.01)
Mark-to-market LTV spline	<=100%	0.641***	0.634***
		(0.029)	(0.03)
	100~120%	0.636***	0.629***
		(0.027)	(0.028)
	120~140%	0.608***	0.602***
		(0.025)	(0.026)
	140~160%	0.577***	0.572***
		(0.025)	(0.025)
	>160%	0.522***	0.517***
		(0.026)	(0.026)
Liquidation timeline spline	<0.5 year	0.071***	0.069***
		(0.011)	(0.011)
	0.5~1 year	0.06***	0.058***
		(0.006)	(0.005)
	1~2 years	0.068***	0.066***
		(0.005)	(0.005)
	2~3 years	0.063***	0.06***
		(0.004)	(0.004)
	3~4 years	0.055***	0.053***
		(0.004)	(0.004)
	4~5 years	0.049***	0.047***
		(0.004)	(0.003)
	>5 years	0.047***	0.044***
		(0.003)	(0.003)
Loans affected by the robo-signing Scandal		0.016***	0.014***
		(0.003)	(0.002)
Loans affected by the National Mortgage Settlement		0.005	0.005
		(0.005)	(0.005)
Loans affected by the CFPB new servicing rules		0.033**	0.042***
		(0.01)	(0.011)
Liquidation volume spline		Y	Y
Liquidation type		Y	Y
Loan and borrower characteristics		Y	Y
Vintage-fixed effect		Y	Y
State-servicer fixed effects		Y	Y
Number of observations		888,593	888,593
Adjusted R-squared		0.4594	0.4621

Notes: Based on the GSE, CoreLogic, and census data. These are generalized least squares (GLS) estimates of the second stage of the Heckman two-stage loss severity regression (equation 3). The dependent variable is the loss severity rate. Loan and borrower characteristics included are mortgage coupon rate, borrower credit score, debt-to-income (DTI) ratio, property occupancy type, and loan purpose. Log default UPB is used as the weight in the regression. Corrected standard errors are in parentheses assuming error-term clustering at the state level. All vintage-fixed effects, except for 1999 (relative to 2016) are insignificant in Model 2; *** for p<0.01%, ** for p<0.1%, and * for p<5%. See Appendix Table 1 for the full model estimates. The following filters are applied to the loan sample: Repurchased loans are excluded; origination LTV ratio has to be between 47 and 100 percent; contemporaneous (mark-to-market) LTV needs to be populated; and, finally, loss severity or foreclosure timeline outliers are excluded. A small number of observations are lost because of the filtering.

Table 4: Loss Severity Simulation Results

	Number of Loans	Actual Sev.	Pred. Sev.	Delta Shorter Timeline	Delta Longer Timeline		Number of Loans	Actual Sev.	Pred. Sev.	Delta Shorter Timeline	Delta Longer Timeline
AK	681	0.252	0.240	-0.068	0.051	MT	1,622	0.290	0.264	-0.059	0.041
AL	13,640	0.373	0.369	-0.071	0.060	NC	23,176	0.333	0.330	-0.064	0.055
AR	5,129	0.363	0.348	-0.067	0.057	ND	299	0.290	0.278	-0.053	0.042
AZ	50,765	0.448	0.447	-0.081	0.041	NE	2,995	0.305	0.296	-0.064	0.061
CA	79,586	0.402	0.401	-0.068	0.035	NH	4,578	0.381	0.383	-0.071	0.059
CO	13,955	0.243	0.238	-0.066	0.051	NJ	16,866	0.474	0.483	-0.051	0.044
CT	7,337	0.437	0.439	-0.063	0.054	NM	4,110	0.345	0.326	-0.055	0.042
DC	392	0.294	0.302	-0.050	0.039	NV	23,393	0.473	0.471	-0.073	0.029
DE	2,215	0.371	0.383	-0.062	0.049	NY	14,880	0.446	0.451	-0.047	0.040
FL	112,860	0.488	0.487	-0.049	0.026	OH	47,582	0.541	0.567	-0.065	0.053
GA	44,033	0.413	0.413	-0.077	0.058	OK	6,618	0.346	0.337	-0.060	0.052
HI	1,340	0.267	0.278	-0.047	0.036	OR	11,660	0.303	0.290	-0.060	0.037
IA	6,767	0.407	0.411	-0.061	0.049	PA	21,989	0.455	0.470	-0.060	0.047
ID	6,263	0.350	0.324	-0.070	0.040	RI	3,410	0.459	0.464	-0.067	0.054
IL	49,328	0.529	0.539	-0.058	0.042	SC	13,918	0.383	0.391	-0.061	0.050
IN	25,170	0.477	0.502	-0.064	0.052	SD	854	0.270	0.248	-0.057	0.042
KS	6,273	0.346	0.348	-0.064	0.052	TN	15,005	0.343	0.339	-0.065	0.058
KY	9,395	0.391	0.403	-0.061	0.050	TX	29,006	0.256	0.234	-0.060	0.052
LA	6,354	0.366	0.363	-0.057	0.048	UT	8,342	0.267	0.247	-0.069	0.045
MA	12,058	0.356	0.359	-0.057	0.043	VA	14,729	0.321	0.319	-0.073	0.058
MD	13,721	0.413	0.430	-0.057	0.045	VT	862	0.461	0.462	-0.051	0.040
ME	2,914	0.460	0.459	-0.054	0.042	WA	18,956	0.314	0.306	-0.059	0.036
MI	71,359	0.531	0.556	-0.075	0.051	WI	17,870	0.442	0.449	-0.066	0.050
MN	23,281	0.382	0.386	-0.069	0.046	WV	3,017	0.426	0.414	-0.081	0.062
MO	21,210	0.409	0.416	-0.071	0.062	WY	794	0.278	0.242	-0.069	0.044
MS	5,663	0.424	0.407	-0.070	0.052	US	898,220	0.425	0.427	-0.064	0.043

Notes: Based on the GSE, CoreLogic, and census data. Default balance weighted average loss severity for each state and for the U.S. is based on loan-level simulations using the loss severity model shown in Table 3. For each loan, in addition to the actual loss severity and model fitted loss severity, we simulate the loss severities for two scenarios, where the liquidation timeline is one year shorter than the actual timeline, and the liquidation timeline is one year longer than the actual timeline. Other covariates in the model such as contemporaneous LTV ratio are recalculated based on the specific scenario.

Table 5: MLE Estimates of the Accelerated Failure Time Model for Liquidation Timelines

	Judicial State	Non-judicial State
Default timing		
After Jan. 2014	1.147***	1.580***
	(0.011)	(0.011)
Oct. 2012–Jan. 2014	1.073***	1.484***
	(0.009)	(0.01)
Sept. 2010–Oct. 2012	1.024***	1.312***
	(0.008)	(0.009)
Nov. 2008–Aug. 2010	0.809***	1.207***
	(0.009)	(0.01)
Feb. 2007–Oct. 2008	0.594***	1.059***
	(0.009)	(0.011)
MLTV	Y	Y
State-level foreclosure pipeline volume	Y	Y
Previous 12-month HPA	Y	Y
Deficiency judgment	Y	Y
Redemption state	Y	Y
Loan characteristics	Y	Y
Borrower credit score	Y	Y
Intercept	Y	Y
Scale	Y	Y
Number of observations	225,244	285,291
-2LogL	439,341	676,186

Notes: Based on the GSE, CoreLogic, and census data. These are maximum likelihood estimation (MLE) estimates of the accelerated failure time model (equation 4). See Cordell et al. (2015) for a more detailed discussion of the model; standard errors are in parenthesis; *** for p<0.01%, ** for p<0.1%, and * for p<5%. See Appendix Table 2 for the full model estimates.

Table 6: Liquidation Timelines Difference-in-Differences Tests

Panel A: Owner-occupied property loans as treatment group

	NMS	CFPB Servicing Rules
Owner	-0.008	-0.028
	(0.014)	(0.016)
Post-event	0.278***	0.336***
	(0.030)	(0.030)
Owner × Post-event	0.264***	0.231***
	(0.030)	(0.030)
Control variables	Y	Y
State × servicer FE	Y	Y
Observations	37,423	55,928
-2LogL	72,744	61,320

Panel B: Loans serviced by the Big 5 Servicers as treatment group

	NMS	CFPB Servicing Rule
Big 5 Servicer	-0.379***	0.143***
	(0.010)	(0.011)
Post-event	0.041**	0.625***
	(0.014)	(0.015)
Big 5 servicer × Post-event	0.136***	-0.118***
	(0.020)	(0.018)
Control variables	Y	Y
State × servicer FE	Y	Y
Observations	87,130	51,485
-2LogL	185,136	54,609

Notes: Based on the GSE, CoreLogic, and Census Data. These panels present the difference-in-differences (DID) tests of the impact of changes in the servicing industry on foreclosure timelines (equation 5). The DID test is in the form of $Y = \beta_1 T + \beta_2 T * A + \beta_3 A + Z'\gamma$, where T represents the treatment group, A represents the period after an event, and the Z vector represents a vector of control variables. The model estimated is an accelerated failure time model as in Table 4. In Panel A, we limit loans to be those serviced by the Big 5 Servicers in judicial states; the treatment group here are owner-occupied property loans, which are the target of both the NMS and the CFPB servicing rules, and the control group are investor property loans, which are not the target of either the NMS or CFPB. In Panel B, we limit loans to be owner-occupied property loans in judicial states; the treatment group is made up of loans serviced by the Big 5 Servicers, which are the target of the NMS. In both panels, the events are NMS in column 1 and CFPB in column 2, respectively. ***, **, and * indicate 0.1%, 1%, and 5% significance, respectively.

Table 7: Liquidation Timelines Placebo Tests

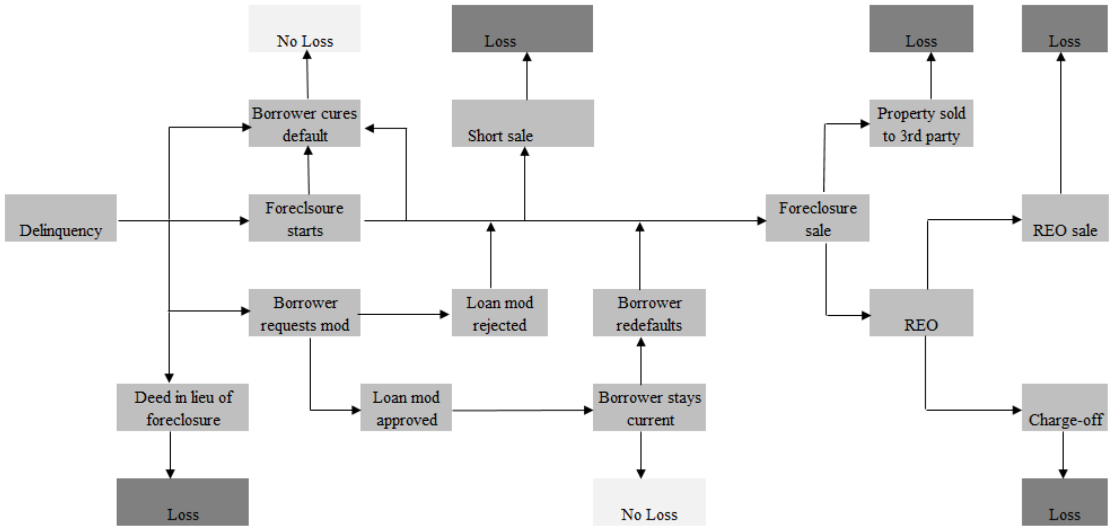
Panel A: Owner-occupied property loans as treatment group

	Foreclosure Moratoria
Owner	0.045
	(0.024)
Post-event	0.660***
	(0.025)
Owner × Post-event	-0.023
	(0.026)
Control variables	Y
State × servicer FE	Y
Observations	25,200
-2LogL	32,831

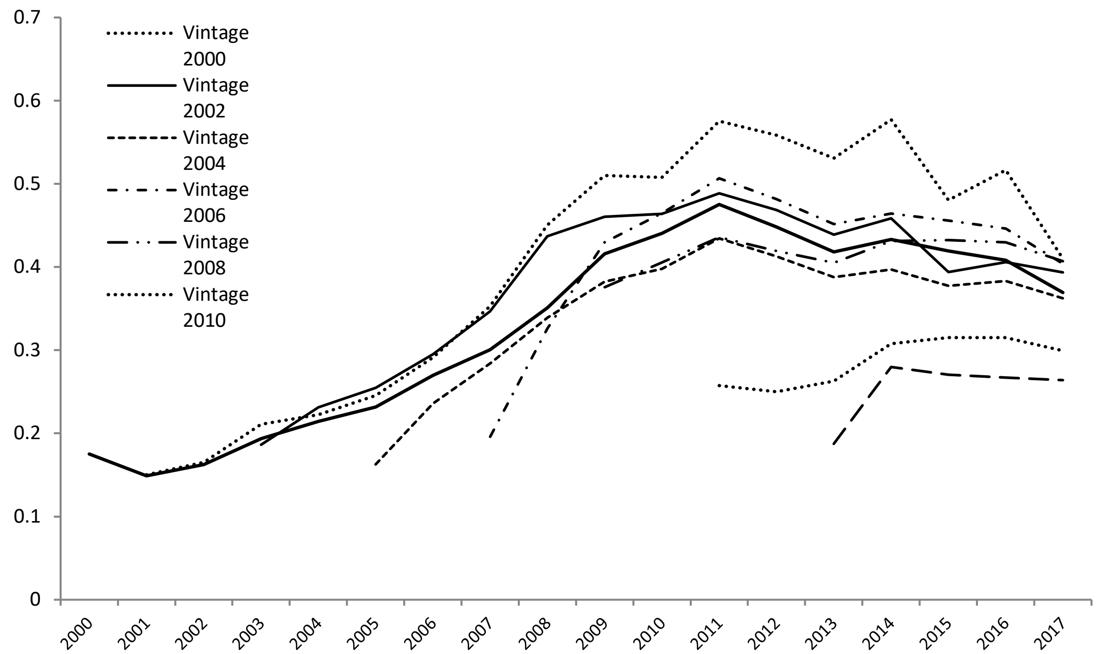
Panel B: Loans serviced by the Big 5 Servicers as treatment group

	Foreclosure Moratoria
Big 5 Servicer	0.054***
	(0.011)
Post-event	0.664***
	(0.010)
Big 5 Servicer × Post-event	-0.019
	(0.012)
Control variables	Y
State × servicer FE	Y
Observations	34,747
-2LogL	47,419

Notes: Based on the GSE, CoreLogic, and census data. These panels present results of placebo tests in a difference-in-differences (DID) framework. The DID test is in the form of Equation 3. The model estimated is an accelerated failure time model as in Table 4. In Panel A, we limit loans to be those serviced by the Big 5 Servicers in judicial states; the treatment group is owner-occupied properties and the control group is investor properties. In Panel B, we limit loans to be those owner-occupied property loans in judicial states; the treatment group is loans serviced by the Big 5 Servicers. In both panels, the event is GSE foreclosure moratoria starting in November 2008. Because the moratoria apply to all loans, we would not expect a significant treatment group effect. ***, **, and * indicate 0.1%, 1%, and 5% significance, respectively.

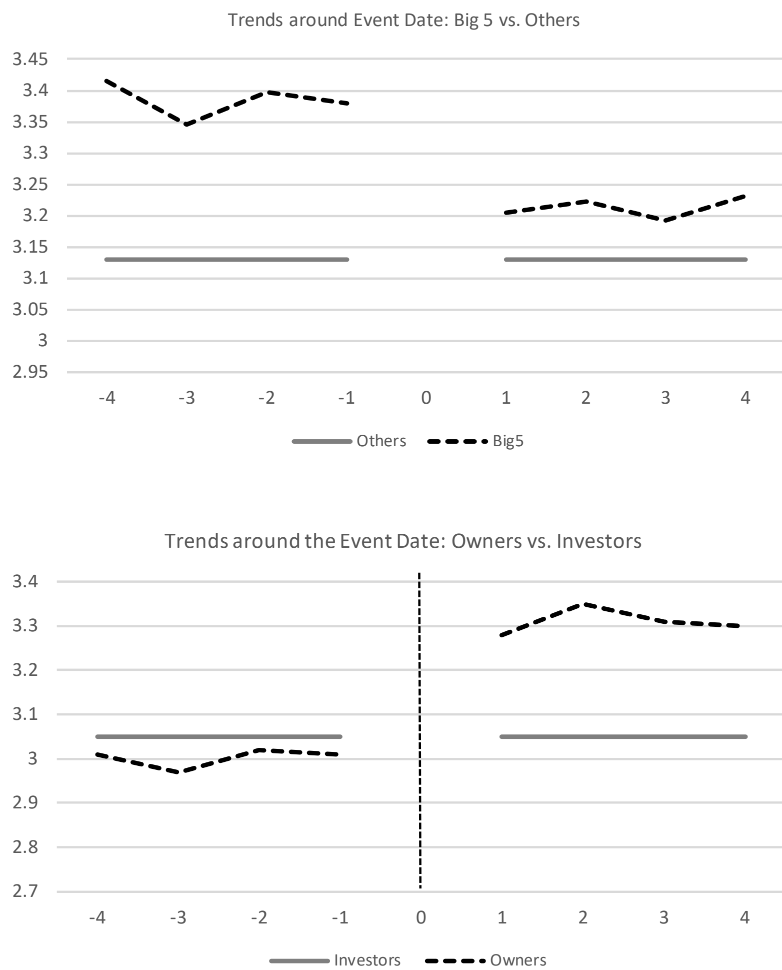


Appendix Figure 1: Typical Default Process



Appendix Figure 2: Loss Severities by Year of Liquidation: By Vintage and Combined

Note: Authors' calculation based on the GSE data



Appendix Figure 3: DID Parallel Trend Test

Notes: These two charts plot results of parallel trend tests. Specifically, in regressions similar to that in equation (5), we interact the treatment variables with time dummies to show how the trends progress period by period. The charts plot the coefficients four quarters before — and four quarters after — the CFPB policy change (event date). Based on the GSE data.

Appendix Table 1: Correlation Between Loss Severity and Selected Covariates in the First Stage of the Heckman Two-Stage Model

	Coupon Rate	Credit Score	DTI	Loan Age	CLTV	10-Year Treasury Rate at Default	GSE Dummy
Loss Severity	0.171	-0.058	-0.048	-0.051	-0.015	-0.014	-0.014

Notes: Based on the GSE, CoreLogic, and census data. In the first stage of the Heckman two-stage model, we include a vector of loan characteristics including mortgage note rate, borrower credit score, debt-to-income (DTI) ratio), and a number of exclusion covariates that help satisfy exclusion restrictions. Our main exclusion covariates include combined loan-to-value (CLTV) ratio, 10-year Treasury rate at loan default, a dummy variable for one of the GSEs, and origination year dummies. Default year dummies, state foreclosure law dummy (judicial vs. nonjudicial) and servicer fixed effect are also included in the first stage.

Appendix Table 2: Full Loss Severity Model Estimates

Covariate		Estimate (S.E.)
Inverse Mills ratio		-0.063***
		(0.01)
Contemporaneous LTV spline	<=100%	0.634***
		(0.03)
	100~120%	0.629***
		(0.028)
	120~140%	0.602***
		(0.026)
	140~160%	0.572***
		(0.025)
	>160%	0.517***
		(0.026)
Liquidation timeline spline	<0.5 year	0.069***
		(0.011)
	0.5~1 year	0.058***
		(0.005)
	1~2 years	0.066***
		(0.005)
	2~3 years	0.06***
		(0.004)
	3~4 years	0.053***
		(0.004)
	4~5 years	0.047***
		(0.003)
	>5 years	0.044***
		(0.003)
Loans affected by the robo-signing Scandal		0.014***
		(0.002)
Loans affected by the National Mortgage Settlement		0.005
		(0.005)
Loans affected by the CFPB new servicing rules		0.042***
		(0.011)
Liquidation volume spline	<1st quartile	-0.012
		(0.078)
	1st~3rd quartile	0.093*
		(0.043)
	>3rd quartile	0.038
		(0.02)

To be continued

Continued from last page		
Default unpaid principal balance	< 30k	0.529***
		(0.022)
	30~ 50k	-0.15***
		(0.01)
	50~100k	0.383***
		(0.017)
	150~200k	0.165***
		(0.01)
	200~250k	-0.074***
		(0.005)
	250~300k	-0.11***
		(0.007)
	>300k	-0.133***
		(0.008)
REO liquidation		0.039***
		(0.008)
Cash-out refinance loan		0.082***
		(0.005)
Rate/term refinance loan		0.065***
		(0.004)
Second home		0.046***
		(0.006)
Investment property		0.101***
		(0.012)
Mortgage coupon rate		0.037***
		(0.003)
Credit score <620		0.015***
		(0.002)
Credit score >720		-0.018***
		(0.003)
Vintage-fixed effects		Y
State servicer fixed effects		Y
Number of observations		888,593
Adjusted R-squared		0.4621

Notes: Based on the GSE, CoreLogic, and census data. These are generalized least square (GLS) estimates of second stage of the Heckman two-stage loss severity regression (equation 3). The dependent variable is the loss severity rate; log default unpaid principal balance (UPB) is used as the weight in the regression; corrected standard errors are in parentheses assuming error-term clustering at the state level; *** for p<0.01%, ** for p<0.1%, and * for p<5%. The reference groups for the categorical variables are default balance between $100k and $150k; home purchase loans; owner-occupied properties; credit score scores between 620 and 720; and foreclosure alternatives (non-real-estate owned (REO)), respectively. See Table 1 for variable definitions.

Appendix Table 3: Full Timeline Model Estimates

	Judicial States	Non-judicial States
Default timing bucket		
After Jan. 2014	1.147***	1.58***
	(0.011)	(0.011)
Oct. 2012–Jan. 2014	1.073***	1.484***
	(0.009)	(0.01)
Sept. 2010–Oct. 2012	1.024***	1.312***
	(0.008)	(0.009)
Nov. 2008–Aug. 2010	0.809***	1.207***
	(0.009)	(0.01)
Feb. 2007–Oct. 2008	0.594***	1.059***
	(0.009)	(0.011)
Contemporaneous LTV Bucket		
<=80%	0.48***	0.471***
	(0.009)	(0.008)
80%–100%	0.308***	0.312***
	(0.008)	(0.007)
100%–120%	0.207***	0.19***
	(0.009)	(0.007)
State-level foreclosure pipeline volume	2.998***	3.557***
	(0.064)	(0.105)
Previous 12-month HPA bucket		
<=-10%	-0.718***	-1.207***
	(0.013)	(0.01)
-10%–0%	-0.446***	-0.947***
	(0.006)	(0.006)
0%–5%	0.007	-0.133***
	(0.006)	(0.005)
Deficiency judgment allowed	0.22***	-0.052***
	(0.059)	(0.005)
Redemption state	0.192***	-0.27***
	(0.008)	(0.005)
Loan size bucket	-0.576***	-0.528***
<= 100k	(0.025)	(0.02)
	-0.383***	-0.311***
100–200k	(0.025)	(0.019)
	-0.159***	-0.134***
200–400k	(0.025)	(0.019)

Continued

Continued from previous page		
Cash-out refinance	0.037***	0.111***
	(0.005)	(0.005)
Rate/term refinance	-0.04***	0.045***
	(0.005)	(0.005)
Owner occupied	0.163***	0.229***
	(0.008)	(0.008)
Second home	-0.005	0.041**
	(0.015)	(0.015)
Borrower credit score bucket		
<=680	0.214***	0.264***
	(0.005)	(0.005)
680–720	0.121***	0.134***
	(0.006)	(0.006)
Intercept	1.76***	1.333***
	(0.065)	(0.024)
Scale	0.847***	1.004***
	(0.002)	(0.002)
Number of observations	225,244	285,291
-2LogL	439,341	676,186

Notes: Based on the GSE and McDash data. These are maximum likelihood estimation (MLE) estimates of the accelerated failure time model (equation 4). See, Cordell et al. (2015) for a more detailed discussion of the model; standard errors are in parenthesis; *** for p<0.01%, ** for p<0.1%, and * for p<5%. The reference groups for the categorical variables are default before February 2007, contemporaneous LTV ratio greater than 120%, previous 12-month house price appreciation greater than 5%, non-deficiency judgment state, non-redemption state, loan size greater than $400k, home purchase loans, investment properties, and credit score greater than 720, respectively.